PAPER VERSUS CRT:

THE EFFECTS OF THE AMOUNT OF TEXT DISPLAYED AND PRESENTATION FORMAT

Thesis

Submitted to

The Arts and Science College of the

UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for

The Degree

Master of Arts in Psychology

by

Benjamin Philip Blinn

UNIVERSITY OF DAYTON

Dayton, Ohio

March, 1998

APPROVED BY:

Dr. David W. Biers, Ph.D.
(Chairperson, Thesis Committee)

Dr. William Moroney, Ph.D.
(Thesis Committee Member)

Dr. Donald Polzella, Ph.D.
(Thesis Committee Member)

Concurrence:

(Dr. F. Thomas Eggemeier, Ph.D.
(Chairperson, Department of Psychology)

ABSTRACT

PAPER VERSUS CRT: THE EFFECTS OF THE AMOUNT OF TEXT DISPLAYED AND PRESENTATION FORMAT

Blinn, Benjamin Philip
University of Dayton, 1998

Advisor: Dr. David W. Biers

The purpose of the present study was to investigate the two factors possibly responsible for differences in searching for information in hardcopy and electronic formats – (1) the amount of text displayed to the subject at any one given point in time and (2) the degree to which the electronic version conforms to a book metaphor. The present study was a 3x2 mixed factorial design in which the presentation mode condition (Hardcopy [paper], Electronic Scrolling, Electronic Page Flipping [book metpahor]) and page length (Short [23 lines] and Long [46 lines]) were varied.

Each subject participated under each presentation mode over three sessions under one of the two page length conditions. During each of the three sessions, the subjects engaged in five different search tasks using the same book under one of the three document mode conditions. The subject was given 10 minutes to complete each search task. If the subject did not find the correct answer within the allotted time, the task was terminated and the next task began. The primary dependent variables were the percentage of tasks completed correctly and four time-based measures -- time to formulate the task,

time to utilize the search tools, time to search the document, and the overall task time. At the end of the third session, the subjects completed a questionnaire dealing with the usability of and user preference for the three document conditions.

The results of the present study showed that there were no significant performance differences in the time it took to formulate the task, utilize the search tools, search through the document, and the overall task time. The only significant performance difference was in the percent of correct measure where the Long Page-Page Flipping document mode resulted in the poorest performance. Despite the failure to find the differences with the performance measures, significant differences were obtained with the subjective data showing a preference for the scrolling mode (electronic format) over the hardcopy mode (paper format). The performance and subjective preference results were most likely due to an interaction of two factors -- the advanced level of computer experience of the subjects in conjunction with the difficulty in utilizing the hardcopy presentation mode.

These results led to three conclusions. First, it was argued that the amount of text displayed hypothesis might have been upheld had less experienced subjects been employed. Second, it was suggested that the locus of the effect was not in the amount of time it took participants to search the document, but rather in the time needed to utilize the search tools. Third, from a design standpoint, one should not use the electronic page flipping -book metaphor unless the page length is limited to the number of lines that can be displayed on the computer screen.

ACKNOWLEGEMENTS

This thesis is dedicated to the memory of my father, Dr. Elliott Blinn. His patient ability to teach, encourage, and motivate helped me to achieve my educational goals. I will always be grateful for his constant guidance, support, and love. I would also like to thank my mother, Mrs. Joyce Blinn, for her continuous support of my academic pursuits.

I would like to express my gratitude and appreciation to Dr. David Biers, my thesis advisor and mentor, for his advice, patience, and direction with my thesis as well as my professional endeavors. I would also like to thank Dr. William Moroney and Dr. Donald Polzella for providing me with their assistance and guidance by serving on my thesis committee.

I would also like to acknowledge Mr. Michael Catani for assisting me with the code for the prototypes using Visual Basic 4.0. I would also like to show appreciation for the thirty-two graduate students in the Department of Psychology for their time and effort in assisting me with my thesis by becoming participants in this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF ABBREVIATIONS

HC.....Hardcopy
ES.....Electronic Scrolling
PF.....Electronic Page Flipping
SP.....Short Pages
LP.....Long Pages
MS.....Microsoft

CHAPTER I

INTRODUCTION

In the library of the future, books will be presented on the computer. The move to the electronic format is the result of a library not being able to store books in available space and the increased cost of printing Hardcopy (HC) books. With libraries and bookstores placing hard copy book information in electronic format, an attempt must be made to understand the difference in the processing of information in its HC and electronic formats.

The present study seeks to understand why information may be found faster in hardcopy format (paper) as opposed to the electronic format of the same document. The focus of the study is on the differences in the amount of text displayed to the user at any one point in time and the failure of the electronic format to the use the book metaphor.

Impetus for the Present Study

The impetus for the present experiment was a combined usability and marketing assessment of a new electronic library system conducted by Biers et al. (1995). The purpose of that study was to determine the value of a product to the end-user and to identify problems in the use of that product. To that end, Biers et al. conducted a comparative usability test in which beginner and intermediate computer users utilized a

hardcopy (paper) book, an electronic version of the same book and WinHelp (a Windows 95 onscreen help guide) to find information about Windows 95. The users engaged in a series of tasks in which they utilized each of the three tools over three sessions. It was expected that the electronic search tool would aid more readily in finding the information because of the sophistication of the search tool. The major finding was that there were no differences between the electronic and HC versions of the book in terms of percent correct. Surprisingly, however, it took the users longer to find the information with the electronic book than the HC documentation. This unexpected finding forms the basis for the present study.

To understand why there were differences between the HC and the electronic versions of the book, it is necessary to decompose the task of searching for information into two stages. In the first stage, the subjects utilized some search tool (i.e. table of contents, index, hypertext links, etc.), to find the approximate location of the relevant information. During this stage, the user must first translate the problem or question into a set of terms and map it onto the terms in the table of contents or index. Then the user must search the index or table of contents to find the target terms. It would be expected that if the user's terms are in the index, then he/she should get to the appropriate part of the book faster using the electronic search tool than thumbing through a book index. But from another perspective, if the user does not immediately find the term he/she is looking for in the index, the electronic index does not give the user a gestalt of possible alternative search terms or entry points. Only a small number of terms appear in the search window at any one point in time. However, with the HC index, many more search terms can be readily scanned to get a broad perspective of possible search terms because

more terms appear in front of the user at any one point in time. Finally, the last step of the first stage involves going to the location where they think the relevant information can be found. The electronic search tool should take the users to that information faster than thumbing to the appropriate page.

In the second stage, the users must scan and search the text of the document to find the answer to the question. This information may be located at the top of the page or across multiple pages. If the information is not found, the users go back to the search tool and select another topic and repeat the process until they find an answer.

In analyzing why there were differences in the task completion for the electronic versus the HC book in the Biers et al. study, several observations were made relative to the two stage model. First, some of the differences could be attributed to the first stage in that users had trouble utilizing the electronic search tool, particularly beginning users.

Another source of difficulty can be traced to the fact that there were differences in the second stage, even when the search tool took them to the correct location. When the subject used the electronic book they seemed to exhibit "tunnel vision" in comparison to the HC book. Namely, there was a reluctance to stray too far from where the search tool sent them within the document. The users tended to stay right on the screen that was in front of them and there was little attempt to scroll beyond what was already there. This was a particular problem when the answer was spread across multiple pages (i.e. screens) or when the answer was at the end of the section. In this situation, the subjects would have had to scan several "pages" (i.e. screens) to find the information. It was as if the users expected the search tool to send them to the exact location of the answers right away. If it did not, the users were reluctant to scan further.

Another possible reason that subjects found the information in the HC documentation faster was that more information was available to the user at any one point in time during the second stage. With the HC documentation, the book has two pages visible, and the number of lines shown at any one page in the book is more than what can fit on a computer screen. The person could readily scan more information, since the book was back-to-back, 2 pages at a time. However, with the electronic book, there is a limitation to the number of lines that are shown on the screen. Also, as the user scrolls down, what he/she has initially seen, disappears.

Finally, the electronic book-HC difference may be a result of the electronic version not conforming to the book metaphor with which users have been familiar since an early age (i.e. well learned habits). In the HC documentation, there are three essential components to this book metaphor. First, the HC documentation is organized in pages, side by side, and one scans for information page to page, horizontally from left to right. Secondly, the text is in a fixed location. As a consequence of this fixed location, the user's eyes move as he/she scans for information from top to bottom. Thirdly, user can tell where he/she is by the thickness of the book.

However with electronic documentation, the information is typically not organized in discrete pages and information moves on the screen as one moves the vertical scrollbar. In the electronic version of the book, the text actually moves upward as the user scrolls down. The user may start out by scanning from top to bottom but when he/she gets to the bottom of the page and uses the scrollbar, the gaze then becomes fixed as the text moves. This text movement could lead to eye fatigue. The lack of page

numbers and the fact that the text does not appear in a fixed location could only add to the user's confusion.

Review of the Literature on Paper-CRT Differences

Dillon (1992) conducted a comprehensive review of the factors that could account for the differences in processing information with hard copy documentation (paper) as opposed to the electronic version (CRT) of the same document. Table 1 presents an overview of the variables researched and the major outcomes. Inspection of Table 1 reveals that much of this research occurred prior to 1992. A personal communication with Dillon (email) reveals that little empirical research on paper-CRT differences has appeared in the open literature since publication of his book.

The majority of work on CRT-paper differences has been conducted by Gould and his colleagues. In most of these studies, they used a proofreading task in which the subject was asked to locate errors in a document. The primary dependent variables were reading speed and accuracy. The initial study by Gould et al.(1984), demonstrated that reading speed was faster on paper than on the computer screen. The magnitude of the difference between the two media was approximately 20-30 percent.

In an attempt to determine the factors responsible for the paper-CRT difference, Gould et al. (1987a) conducted a series of experiments in which they varied a number display, task, and personal variables—display orientation, visual angle, reading distance, eye movements, display contrast ratio, polarity, font, and user experience. In each of these experiments, Gould (1987a) found there was a reading speed difference with subjects taking longer to find the proofreading errors using the computer screen than the

Table 1

Literature Review

Factor	Study	Variables	Results- Search Times	Results- Accuracy	Results- Other DVs	Conclusion
Same work using CRT and Paper	Gould (1984)	Paper (P) vs. CRT (CRT)	P<CRT			Subjects read from paper 20-30% faster than CRT
Orientation - Paper is long and narrow (vertical) where CRT is short and wide (horizontal)	Gould (1987a)	CRT, Paper Vertical (PV), paper Horizontal (PH)	PH=PV<CRT	PH=PV=CRT		Orientation does not account for P-CRT difference
Visual Angle (width of text line) -CRT display generally has wider text lines than paper	Gould et. al.(1987a)	Paper(P) vs CRT images as a function of visual angle of text in degrees (6.7, 10.6, 24.3, 36.4, 53.4)	6.7: P<CRT 10.6: P<CRT 24.3: P<CRT 36.4: P<CRT 53.7: P<CRT Diff btn P and CRT decreases as fcn of visual angle	6.7:P>CRT 10.6: P>CRT 24.3: P=CRT 36.4: P=CRT 53.7: P=CRT		Some of P-CRT diff could be due to visual angle--interaction suggests this
Orientation and Task - Proofreading v. Comprehension	Gould et. al. (1987a)	Proofreading(Prf) v Comprehension(C) as fcn of Paper Vertical (PV), Paper Horizontal (PH), or CRT	Prf:PH=PV<CRT C:PV=PH C:PV<CRT C:PH=CRT	Prf:PH=PV=CRT C:PV=PH=CRT		Some of diff could be due to orientation for comprehension task
Reading Distance	Gould et. al. (1987a)	Reading Distance (cm) as fcn of Paper (P) v CRT			Read Dist: CRT>P	Avg. visual angle less for paper-visual angle could affect P-CRT diff
Eye Movements	Gould et. al. (1987a)	CRT v Paper (P)			Eye Fixations per line:P<CRT Pattern of Fixations same	Pattern of Fixtrions same indicates that subjects did not tune out, change their scan behavior, or lose places with CRT
Dynamic Characteristics of CRT	Gould et. al. (1987a)	Paper (P) v CRT(Dynamic) v Photos of CRT screen (NonDynamic) (CRT-P)	P<CRT=CRT-P	P=CRT-P>CRT		No evidence that dynamic characteristics of CRT explained P-CRT diff

Table 1 (continued)

Factor	Study	Variables	Results- Search Times	Results- Accuracy	Results- Other DVs	Conclusion
General Quality of CRT Displays - Good v Bad	Gould et. al. (1987a)	Paper (P) v IBM 3278 (3278) v IBM 3290 (3290) v 3277 (3277)	P<3278=3290= 3277			Display quality did not affect the P-CRT difference
Display Contrast Ratio of Text v Background	Gould et. al. (1987a)	Display Contrast ratio of Paper(P) vs CRT			Contrast Ratio:P=CR T	Display contrast ratio did not affect the P-CRT difference
Polarity(dark char on light bkgrnd v light char on dark bkgrnd) and Aspect ratio	Gould et. al. (1987a)	CRT vs CRT screen photos as fcn of Polarity(dark char on light bkgrnd (D/L) v light char on dark bkgrnd (L/D)) and aspect ratio (CRT (AR/CRT)vs Paper (AR/P)	CRT=L/D-AR/CRT =L/D-AR/P =D/L-AR/CRT =D/L-AR/P	CRT=L/D-AR/CRT =L/D-AR/P =D/L-AR/CRT =D/L-AR/P		Neither Polarity nor Aspect Ratio affected P-CRT diff
Font- In previous studies, font differed btn paper and CRT	Gould et. al. (1987a)	Paper(P) vs. CRT with same font	P<CRT	P=CRT		P=CRT when use same font but the polarity differed (not controlled)
Experience in Reading from CRTs	Gould et. al. (1987a)	Paper(P) v CRT as fcn of Experience(Inexperienced vs. Experienced)	I: P<CRT E: P<CRT			Experience does not affect P-CRT diff
Font, Polarity, Size, and Color- First exp in which all these factors were identical in Paper(P) and CRT	Gould et. al. (1987b)	Paper (P) v CRT as fcn of three different fonts	P<CRT for all 3 fonts but P-CRT diff reduced to 5%	P=CRT for all 3 fonts	Preference: P>CRT	With font, polarity, size and color controlled P-CRT diff still signif but reduced
Page Length and Higher Quality CRT-Font, Polarity, Size, and Color were identical in Paper(P) and CRT but with higher quality CRT display (50 hz refresh rate)	Gould et. al. (1987b)	Paper (P) v CRT as fcn of page length (22 lines v 28 lines per page)	22 lines:P=CRT 28 lines: P=CRT	22 lines:P=CRT 28 lines: P=CRT		With higher quality CRT display P=CRT and number of lines per page did not affect this result.

Table 1 (continued)

Factor	Study	Variables	Results- Search Times	Results- Accuracy	Results- Other DVs	Conclusion
Page Length and Higher Quality CRT - Used CRT with 60 Hz refresh rate to reduce flicker	Gould et. al. (1987b)	Paper (P) v CRT as fcn of page length (22 lines v 28 lines per page)	22 lines:P=CRT 28 lines: P=CRT	22 lines:P=CRT 28 lines: P=CRT		With higher quality CRT display P=CRT and number of lines per page did not affect this result.
Display Polarity	Gould et. al. (1987b)	Paper (P-D/L) v CRT-Light Char on Dark Bkgrnd (CRT-L/D) v CRT Dark Char on Light Bkgrnd (CRT-D/L)	P-D/L = CRT-L/D= CRT-D/L	P-D/L = CRT-L/D= CRT-D/L	Preference: P-D/L>CRT CRT-L/D> CRT-D/L	Since all conditions equal, differences in polarity not a factor for improvement in CRt reading
Anti-Alaising	Gould et. al. (1987b)	Paper v aliased CRT (CRT-A) v anti-aliased CRT (CRT-AA)	P<CRT-A **P = CRT-AA** CRT-A=CRT-AA	P=CRT-A=CRT-AA	Preference: P>CRT-AA>CRT-A	Anti-Aliasing accounts for some of the improvement in CRT reading
Display Quality	Gould et. al. (1987b)	Print Quality (Good Quality Print Paper (GQP) v Poor Quality Print Paper(PQP)) and CRT Quality(High=Anti-Aliasing Display (CRT-AAD) v lower quality= IBM 3278 (CRT-3278)and IBM PC (CRT-PC))	GQP<PQP; GQP<CRT-3278 =CRT-PC; **GQP=CRT-AAD**; CRT-AAD =CRT-PC<CRT-3278	GQP=CRT-AAD=CRT-PC =CRT3278>PQP	Preference: CRT-AAD> GQP >CRT-PC>CRT-3278>PQP	Anti-Aliased Displays can be read as fast as good quality paper
Text Format- Typical of Book (60char/line;40 lines) vs Typical of CRT (39 char/line;20 lines)Line Length and Characters per Line	Kruck & Muter(1984) Experiment 1	Paper-Book format (P-B) v Paper-CRT format (P-CRT) v CRT-CRT format (CRT)	P-B<P-CRT<CRT		Comprehension P-B=P-CRT =CRT	Putting Paper in CRT format reduces the Paper-CRT diff
Scrolling/Paging	Schwarz, Beldie, and Pastoor	CRT: Scrolling (CRT-S) vs Paging (CRT-P)	(CRT-S)=(CRT-P)		Preference: Pag>Scroll	

Table 1 (continued)

Factor	Study	Variables	Results- Search Times	Results- Accuracy	Results- Other DVs	Conclusion
Readability of scrolled CRT text as fcn of Line Length, Number of Char per Line, and Number of Lines	Duchnicky and Kolers (1983)	CRT: Line Length(1/3, 2/3, full); Densities of characters(40 char/line (40C)v 80 char/line(80C,); Number of lines(1L, 2L, 3L, 4L, 20L)	1/3<Full<2/3; 80C<40C; 4L=2L;20L<1L=2L	1/3=Full=2/3; 80C=40C; 1L=2L=3L=4L= 20L		These three variables affected the rate of which the scrolled text was read
Text Splitting across screens	Dillon Richardson, McKnight (1990)	CRT: 60 lines (non-Split (60-NS) and Split Text (60-S) vs. 20 lines (Non Split (20-NS) and Split (20-s)	(60-NS)=(60-S)=(20-S)=(20-NS)		Comprehension:(60-NS)=(60-S)=(20-S)=(20-NS) Preference: 60>20 Lines	The effect of sentence splitting has no effect on comprehension or speed
Effect of Window Size on Reading Electronic Text	Richardson J., Dillon A., McKnight C. (1989) Experiment 1	CRT Page Flipping: Screen Size in Lines 20L, 40L)	20L=40L		Time Spent in TOC and Index as a fcn of Window Size:20L=40L; Preference: 20L<40L	Subjects using the small screen page more than subjects of large screens in order to read same info
Effect of Window Size on Reading Electronic Text (only change of direction of >2 pages)	Richardson J., Dillon A., McKnight C. (1989) Experiment 2	CRT Page Flipping: Screen Size in Lines 20L, 40L)			Comprehension:40L<20L Preference: 20L<40L	Screen size does not affect Comprehension and perf. rates

paper version. No single variable was able to account for the 20-30% difference in reading speed.

Gould et al. (1987b) reported several additional experiments in which higher quality CRT monitors were utilized and which exerted greater control over the quality of the text. In Gould et. al's previous studies, there were a number of factors that were not held constant between the two media. For example, the printers used to print the HC version differed from the CRT display in terms of font, polarity, size, and color. Gould et al.'s (1987b) results indicated that when the display characteristics were identical on the paper and computer screen, the differences in reading speed between the two media were reduced to a mere five percent. With a higher quality CRT (e.g. 60mHz refresh rate with anti-aliasing) this difference disappeared altogether. As a consequence, Gould et al. (1987b) concluded there was no inherent difference in reading speed between the computer screen and paper as long as a high resolution display is used and the font, polarity, color, and size are identical.

This conclusion is in apparent contradiction to the Biers et al. (1995) findings. However, the Biers et al. (1995) study differed in three fundamental ways. First, the Biers et al. (1995) study used a full-length reference book as opposed to a short passage. Second, the electronic (CRT) information was in the form of a scrollable text so the information moved as the subject used the vertical scrollbar. Third, the Biers et al. (1995) study used a combined search and comprehension task whereas Gould focused upon the reading speed associated with proofreading.

Factors relevant to the present study

Several of the studies reviewed by Dillon are particularly germane to the present study. These include scrolling versus paging, display size, and text splitting across screens. These studies are discussed below around the two central variables being manipulated in the present study.

Book Metaphor

One of the potential reasons why electronic and HC differ is the degree of which the electronic format conforms to the book metaphor. The typical electronic text mode involves scrolling vertically through the information whereas the book involves the user flipping through the pages. An attempt has been made by several researchers to make the electronic version look more like a book by breaking the text up into discrete pages.

There are two previously researched factors which related to the book metaphor concept. These factors include paging versus scrolling and text splitting across the screen. For example, Schwarz et al. (1983) manipulated the textual information via scrolling in two different ways. The information could either be scrolled continuously or line by line. In the paging format, the textual information changed one screen at a time. The subject could control the screen content by the use of a joystick. Moving the joystick forward resulted in upward scrolling and backward for downward scrolling. The joystick was also used to change the screen contents in the paging model. Schwarz et al. (1983) hypothesized that if the parts of the answer were in close proximity to one another, scrolling would be better because the subject could view the parts of the answer on the computer screen. Schwarz also hypothesized that paging would be better if the relevant parts of the answer were further apart. Schwarz et al. (1983) found that the mode of

operation (paging versus scrolling) had no effect on reading speed or percentage of correct answers. However, the subjects preferred paging to scrolling.

An important issue is the way that the text is split between screens in a paging format. This issue is particularly germane to computer screens because the information is presented one screen at a time and the break between the screens is likely to be more critical. However, with a book (paper), the subject could view two pages at a time and access to previous pages is generally seen as effortless.

Dillon et al. (1990b) examined the role of text splitting on performance by manipulating a paging mode. Dillon et al. (1990b) used display sizes of 20 and 60 lines with non-split text and split text. The dependent measures included comprehension and reading speed. The experimenters used a lengthy text as well as an older style page manipulation tool. This tool contained the following commands: Next Page, Previous Page, Go to the Beginning, Go to the End, Go to a specific page, and Quit the Program. The results revealed no significant effect for screen size or sentence splitting on either reading comprehension or reading speed.

These results suggest that paging versus scrolling and text splitting does not account for the differences in subject performance while reading information on the computer screen (CRT).

Amount of Text Displayed

A second factor that might account for differences between the two media is the amount of text information in view of the user at any one point in time. One of the explanations given for the Biers et al. (1995) findings was that more information was

displayed at any one point in time in HC documentation as opposed to the electronic book.

One way of circumventing this problem is to increase the number lines that appear on the computer screen. Computer screens typically hold 25 lines and it would seem intuitive that the greater the number of lines (display size), the easier it would be to locate and retrieve the relevant information. A larger screen size should result in the subject manipulating the electronic text (move forward or backward) less frequently because there is more information per screen.

The results of Duchnicky and Kolers (1983) and Elkerton and Williges (1984) seem to contradict that belief. Duchnicky and Kolers (Table1) stated that there was little to be gained by increasing the display size to more than 4 lines. Elkerton and Williges (1984) varied display size using 1,7, 13, and 19 line displays with text that was constantly scrolled. They reported that there were few speed or accuracy advantages between the displays of seven or more lines. However, these studies used passages that were no longer than 300 words and only the Duchnicky and Kolers (1983) experiment used reading comprehension as a dependent measure.

Richardson et al. (1988) attempted to overcome these shortcomings by presenting the subjects tasks that they might encounter in the real world. The first experiment used a software reference book and the second experiment used an academic journal. The subjects were asked to locate specific information within the text in order to answer questions with the software reference book and to read for comprehension using the academic journal. The text was presented as single screens with a Next Page/Previous Page buttons to help with navigation. Richardson et al. (1988) used screen sizes of 20

and 40 lines for the reference manual and 20 and 60 lines for the journal article. For the first experiment, Richardson et al. (1988) found no significant effect for display size on the time to complete the tasks although the large screen resulted in faster search times for seven of the ten tasks. Subjects expressed a preference for larger windows than smaller windows. They expressed their dislike for the constant page-turning with the small windows even when the subject wished to read a small portion of the text. Subjects also complained that they constantly forgot which section or chapter heading they were located when using the smaller window. The results of the second experiment, which used the academic journal, indicated that there was no significant effect for display size although again there was a slight trend favoring the larger screen.

Present Study

The purpose of the present study was to investigate the factors responsible for differences in searching for information in its HC and electronic formats. The two reasons being investigated were the amount of text which is displayed to the user at any one given point in time and the failure of electronic version to conform to the book metaphor.

To investigate these factors, the present study used a 3x2 mixed factorial design in which the presentation mode (HC documentation, Electronic Scrolling [ES], Electronic Page Flipping [PF]) and page length (short and long) were varied. The same federal income tax guide was used in all three document conditions. The users were asked to perform several search tasks utilizing each of the three document conditions over three sessions. One half of the subjects received the short pages (SP) across all three document conditions, and the other half received the long pages (LP) across all three media.

The first variable being investigated was the presentation mode. The HC condition was a facsimile of a real book in which the material was printed on paper presented as single-sided pages, not back-to-back, so the user could see only one page at a time. The other two document conditions involved presenting the same book on the computer with search tools to aid the user in finding information. The ES format, the typical way of displaying a book on the computer, presented the text as a continuous vertical scroll with a vertical scrollbar to aid the user in moving through the entire book. The PF mode presented the information as discrete pages with a button to turn the page. This format allowed the user to see one page at a time. In addition a horizontal scroll bar was available to move or flip through the pages. The reason for designing the PF condition was to make the electronic format more closely resemble the HC version.

The second independent variable was page length, or the amount information that is presented to the user at any one time. A short page was defined as the amount of text that could fit on a computer screen at any given time, 23 lines per page. In the HC-SP condition, the information was presented one page at a time, front-side only, with the same number of lines as contained on a computer screen (23 lines). An LP was defined as the amount of text displayed that is contained in a page in the HC version of the actual book, 46 lines. The LP version of the electronic format contained the same information but the user would have to use the vertical scrollbar in order to view the remainder of information on a page since only 23 lines could be viewed at any given time.

In generating predictions about the expected outcome, it is important to recall the two stage model for searching the information that was proposed earlier. The first phase consists of using the search tool to find and go to the presumed location of the

information. The second stage involves the user finding the correct information once the search tool sends them to the text. If the information was not found on those pages, the user repeats the two stage process.

The present study attempted to minimize search differences which could have resulted as a consequence of the first stage by creating a relatively low level electronic search tool and by controlling the amount of text available to the user at any one point of time in the index. The search tool was low level in that it basically replicated the table of contents and the index of the HC documentation. The electronic search tool had no hyperlinks or word searches to locate specific words in the text or index. The only factor that could make the electronic version better in Stage 1 was that the electronic search automatically took the users to the location that the book selected rather than the user having to thumb through the pages to find the information. This advantage, however, may be counteracted by the fact that the user had to gain experience and familiarity using the search tool. Thus, it was expected that these two factors would work against one another to produce little or no differences in the time it takes to find the information.

Secondly, the amount of text displayed which could be viewed in the index at any one point in time was controlled. In the typical HC book, the user can view and scan more information at any one point in time because more index terms can fit on a paper layout (typically multiple columns and two pages at a time). However, with electronic documentation, the user can only see a fraction of the index and therefore, had to use the vertical scrollbar to view the remaining information. For the purposes of the present study, the number of lines in the index that a user could view at any given time with the

HC and the electronic versions was held constant. That is, the number of lines in the index of the HC format will be equivalent to the electronic index.

The focus of the present research is on the amount of text information that was displayed to the user during Stage 2 and the degree to which the format conforms to the typical book metaphor. The predictions that follow are based upon the assumption that there will be no differences in the use of the search tool (i.e. Stage 1.)

If the amount of text information is a factor that accounts for the differences between the HC and electronic format (i.e. short versus LP's), there will be an interaction of page length and document condition. The underlying hypothesis is that the greater the amount of text which can be viewed at any one time, the faster the search time. With LP's, the HC documentation should provide shorter search times than the electronic book because there is more text information in view of the subject at any one point in time. However, with SP there will be no difference between the HC and ES format mode because the amount of text information displayed to the user will be equivalent.

The second factor being explored is the degree to which the CRT format approximates the typical book. Both the PF format and HC book share a common mental model, that of a typical HC book. However, the ES presentation mode does not conform to this metaphor. Therefore, holding everything else constant, one would expect the PF mode and the HC format to result in faster search times than the ES mode.

More specifically however, the degree to which the book format conforms to the metaphor of a book should interact with page length. Therefore, using the book metaphor as a basis for prediction, there should also be an interaction between page length and document condition, but the nature of the interaction should be different. With SP, the

PF mode should be equivalent to the HC documentation because there is no scrolling within a page under this condition. The information remains within a constant location, and the amount of text displayed to the person is the same. Since these two modes conform to the book metaphor, they should have faster search times than the ES format.

With LP, however, the PF mode no longer conforms to the book metaphor. The user must manipulate the vertical scrollbar to see the remaining information that is hidden. This is an awkward and unfamiliar mode in that the user not only has discrete jumps from page to page but he/she will have to use the vertical scrollbar to view the entire contents of a single page. In this case, the search times should be the longest for the PF mode, intermediate with the Electronic Scrolling (ES), and shortest for the HC condition.

In a comparing the two predictions based upon these alternative two explanations, there are two major differences. First, with SP, the amount of text displayed hypothesis leads to the prediction that there will no differences between the HC documentation and the ES format. However, using the book metaphor explanation, it is predicted that the HC documentation will result in faster search times than the electronic book scrolling mode for SP.

Secondly, for LP, the difference lies in comparing the two electronic versions. Based upon the amount of text displayed prediction, there should be no differences between the two electronic versions since the amount of text which the user can view at any one time is the same. However, using the book metaphor explanation, it is predicted that the ES mode will result in shorter search times than the PF presentation mode because the ES presentation mode does not conform to the book metaphor. The direction

of the outcome of the study is expected to shed some light on which of the two alternative explanations is more cogent.

CHAPTER 2

METHOD

Design

The present study is a 3 x 2 mixed factorial design in which the presentation mode (HC Documentation, ES, PF) and page length (Short and Long) were varied. Subjects were randomly assigned to one of two page length conditions (between-subjects factor). Each subject participated under each presentation mode (the within-subjects factor) over three sessions with the order of the presentation mode conditions being counterbalanced by means of a 3 x 3 Latin square. Under all of the conditions, subjects performed structured search tasks in the presence of an experimenter.

Subjects

Thirty graduate subjects (16 females, 14 males) were recruited from students attending the University of Dayton, a private midwestern university. The subjects were treated in accordance with the ethical Principles of Psychology and the Code of Conduct.

Materials and Apparatus

Sourcebook

The three document conditions were based upon a tax reference document titled, 1996 Your Federal Income Tax for Individuals-Publication #17. A tax reference guide was chosen primarily because it would be motivating and interesting to college students

and also did not require copyright release since it was a government document. The book was a typical reference book which contains a table of contents and a detailed index. The original document was downloaded from the Internet, and saved as a Microsoft Word document. For the purposes of the present study, only 200 pages of the document were utilized. The pictures, diagrams, and tables were eliminated because the Visual Basic custom control text box would accept text only. Once the text of the book was placed into the Word document, it was reformatted and fed into the electronic versions as a file or printed out on paper as a HC documentation so that the font (style, size), number of lines, and width of page were exactly the same in all three document conditions.

Document and Presentation Mode Conditions

Two search tools were provided to subjects in all document conditions--the table of contents and the index. To control for the amount of text that appeared in the index and the table of contents, the information contained in these search tools were identical across all three presentation modes. To control the amount of text that could be viewed by the subject at any point in time, the index was formatted in a single column with 23 lines per page. The format of the index was displayed such that the subtopics under a given main topic were indented. If a search term could be accessed in multiple ways (e.g. on several different pages), the index contained separate lines with the different page numbers.

Hardcopy Documents

The text for Hardcopy-Long Page (HC-LP) condition was printed on paper with 46 lines per page, single-sided, and was placed into a loose leaf binder with the table of contents in front and the index in back. The page numbers appeared at the bottom of the page. This resulted in a long HC book of 206 pages.

The Hardcopy-Short Page (HC-SP) condition was prepared the same way with the exception that there were only 23 lines per page. Thus, the book was approximately twice as thick as the HC-LP format (379 pages).

Electronic Versions

Both electronic versions of the book were developed using the 16-bit version of Microsoft Visual Basic 4.0, to construct a functional interface for the two electronic modes. The interfaces used the standard Windows interface style. The main screen contained a menu bar and tool bar at the top as well as a large text window where the actual text of the book could be viewed. The text window allowed for only 23 lines to be displayed at any one point in time. At the bottom right of the screen, a box contained the page number being viewed. A vertical scrollbar was provided on the right side to move easily through the text.

The menu bar contained the following pulldown menus: File, Edit, Bookmark, Search, Tool Tips, and Help. Clicking on any of the above menu items resulted in a pulldown list that provided additional functions. Table 2 details the pulldown menu for each item on the menu bar and whether or not the item was functional in the mode.

Table 2

Electronic Scrolling and Page Flipping Features

The following features are part of the Electronic Page Flipping mode Menu Bar and whether or not they are disabled for this experiment

- File
  - Open (Disabled)
  - Print (Disabled)

- Edit (Disabled)
- Bookmark (Disabled)
- Search
- Tooltips
- Help (Disabled)

The following features are part of the ES and PF presentation mode Toolbar and whether or not the feature is disabled for the experiment

- Backward (Not present for the ES presentation modes)
- Forward (Not present for the ES presentation modes)
- Index
- Table of Contents
- Add Bookmarks (Disabled)
- Go to Bookmarks (Disabled)
- Print (Disabled)

The remaining features are located on the computer screen for:

- Vertical Scrollbar (Not present for PF-SP)
  Horizontal Scrollbar (Not present for both ES-SP and ES-LP presentation modes)

The toolbar for the ES (Table 3) and PF (Table 4) presentation modes contained an Add Bookmark, Bookmark, Search, Print, and Table of Contents buttons. However, the PF presentation mode presentation mode also contained a backward and forward button.

Table 3

Toolbar (icons and their descriptions) for the Electronic Scrolling Mode

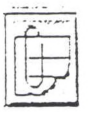
Icon:					
Meaning:	Index	Table of Contents	Add Bookmark	Bookmarks	Print

Table 4

Toolbar (icons and their descriptions) for the Electronic Page Flipping Mode

Icon:							
Meaning:	Back one page	Forwards one page	Index	Table of Contents	Add Bookmark	Bookmarks	Print

The Add Bookmark and Bookmark features were disabled for the present study. When the mouse was placed on a button, a tooltip appeared that displayed the function of the button.

The table of contents (Figure 1) could be accessed in two ways--either through the Search pulldown menu or by selecting the table of contents button. When the table of contents was accessed, a dialog box displayed the 10 chapter titles. There was a plus symbol on the left side of the chapter titles, which when clicked, expanded the table of

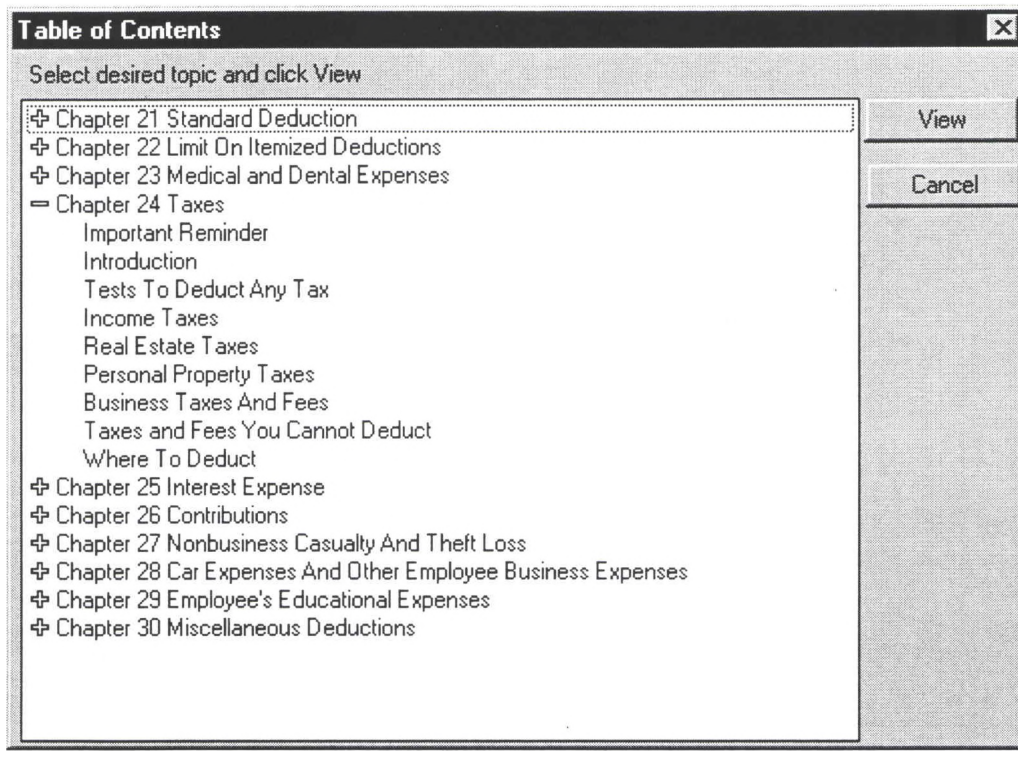


Figure 1. Table of Contents for the Electronic Presentation Modes

contents into its subtopics. The text of a subtopic or chapter could be viewed by double-clicking on the particular item. This action resulted in the appropriate text being displayed in the text window with the first line of the page being at the top of the window.

The index (Figure 2) could also be accessed through either the pulldown menu or the Index button. When the index was accessed, a dialog box appeared which contained: (1) a text box in which the subject typed in a search term and (2) a list box in which the index of topics appeared. The relevant page numbers were included in the list box following the topics and subtopics. The list box allowed 23 lines in the index to be shown at any one point in time. The action of the list box and the text box were coordinated such that when the subject typed in the search term, the search tool jumped to the appropriate section of the index (which is shown in the list box). This action of the index dialog box followed standard windows protocol, namely as each letter was typed, the search tool moved to the part of the index which matched the letter sequence. The list box also contained a vertical scrollbar which allowed the subject to move through the index, independent of having to type information into a text box.

To access a particular topic or subtopic in the index, the subject had to first highlight the item then either double-click on it or hit the 'View' button. At this point, the dialog box disappeared and the text containing the topic appeared in the text window with the top of the page being the first line in the text window.

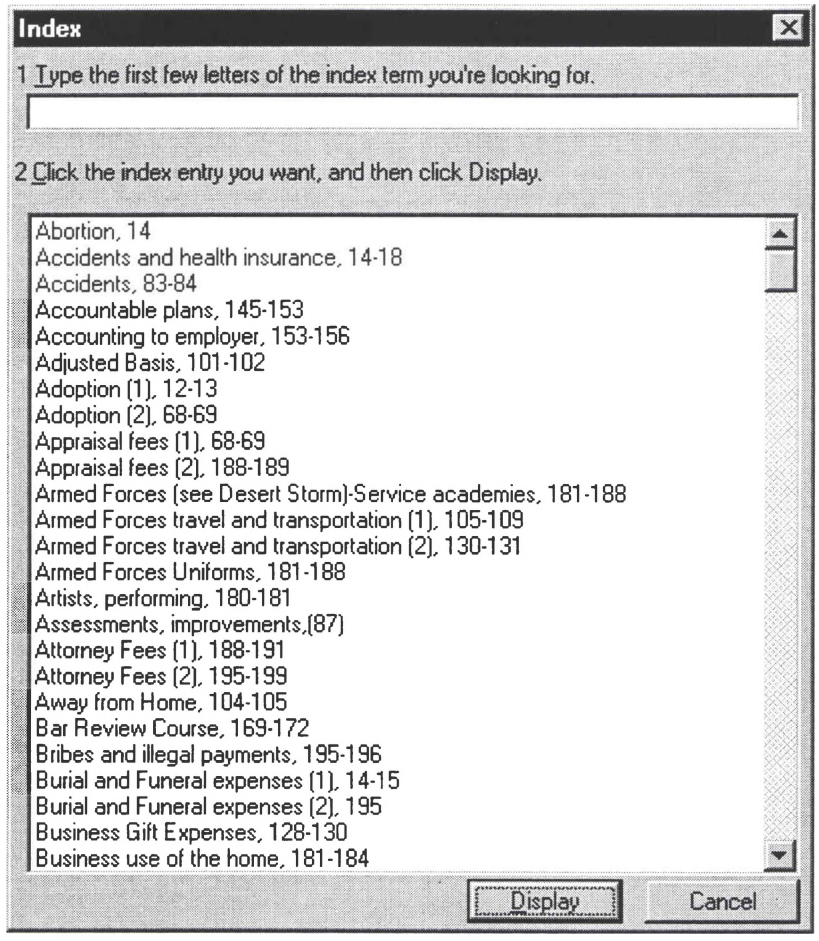


Figure 2. Index of the Electronic Presentation Modes

Electronic Scrolling (ES) Mode. Figure 3 shows an image of the ES mode. Most of the features of this mode have been described earlier. The text of the book covered the largest portion of the screen and permitted only 23 lines to be displayed at any one point in time. To scroll through the entire document, a vertical scrollbar was provided on the right side of the mode. The subject could click on the up arrow (front of the book), the down arrow (back of the book), or drag the vertical scrollbar to the appropriate section the book. In order to view what page was present, the subject manipulated the vertical scrollbar and scrolled through the document until the page number, located in the center of the text followed by a red horizontal line, was present.

There were two files created for the short and LP version of the book. In each file, the book was stored as a single document. To indicate a page break, the text contained a dashed line followed by the page number in the center followed by another dashed line. For the Electronic Scrolling-Long Page condition (ES-LP), the page break occurred every 43 lines. However, in the Electronic Scrolling-Short Page (ES-SP) condition, the page break occurred every 23 lines.

Electronic Page Flipping (PF) Mode. In the PF condition (Figure 4), the mode contained the same information as the ES condition with two exceptions: First, Backward and Forward buttons were available on the toolbar so the subject could move one page at a time. These buttons acted the same way as the Forward and Backward buttons in Netscape. Second, the mode contained a horizontal scrollbar which indicated the relative position within the text. Movement of the horizontal scrollbar to the left moved the page closer to the front of the back, whereas movement of the horizontal scrollbar to the right moved the page closer to the back of the book. The action of the scrollbar was

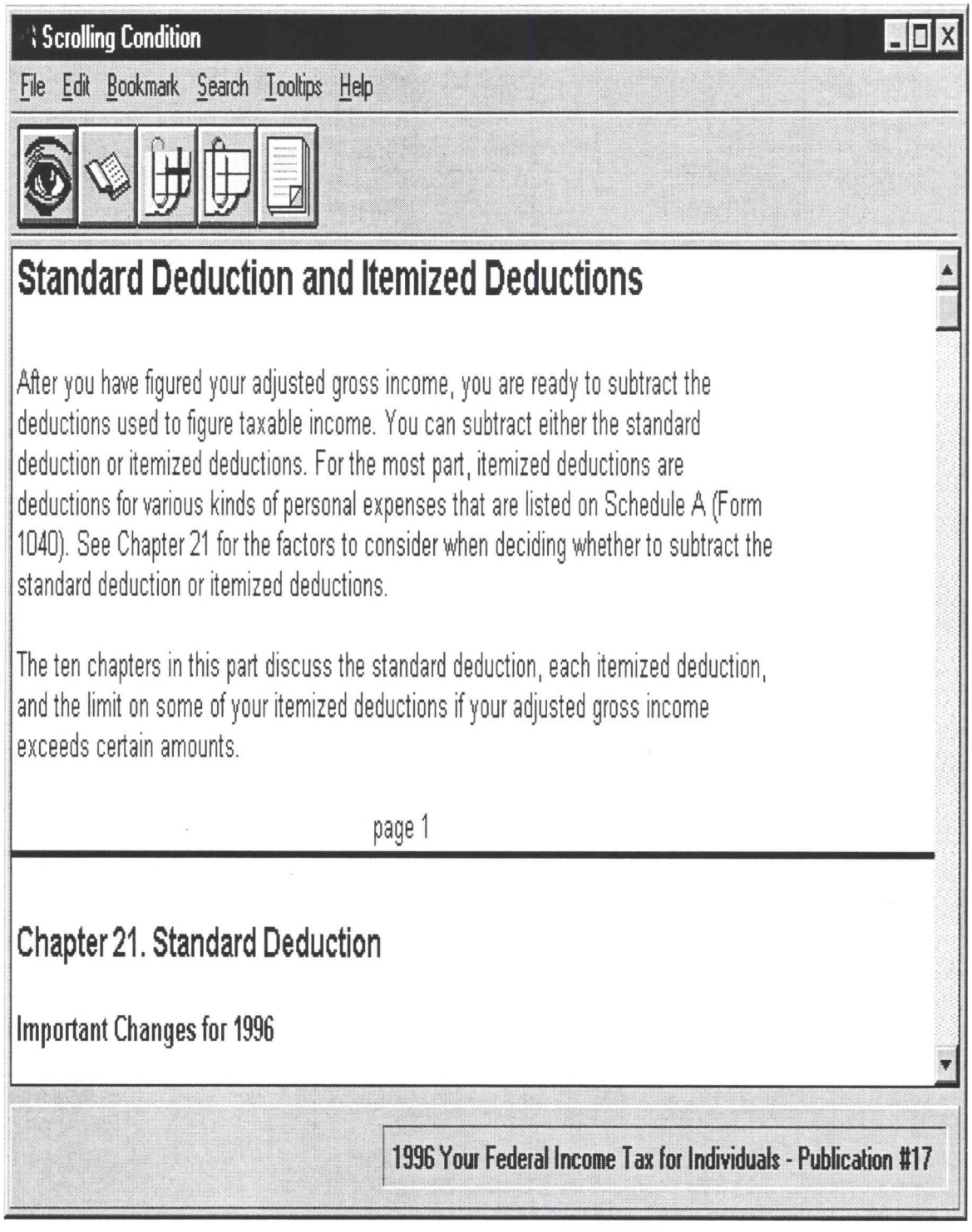


Figure 3. Sample of page as presented in the Electronic Scrolling Mode

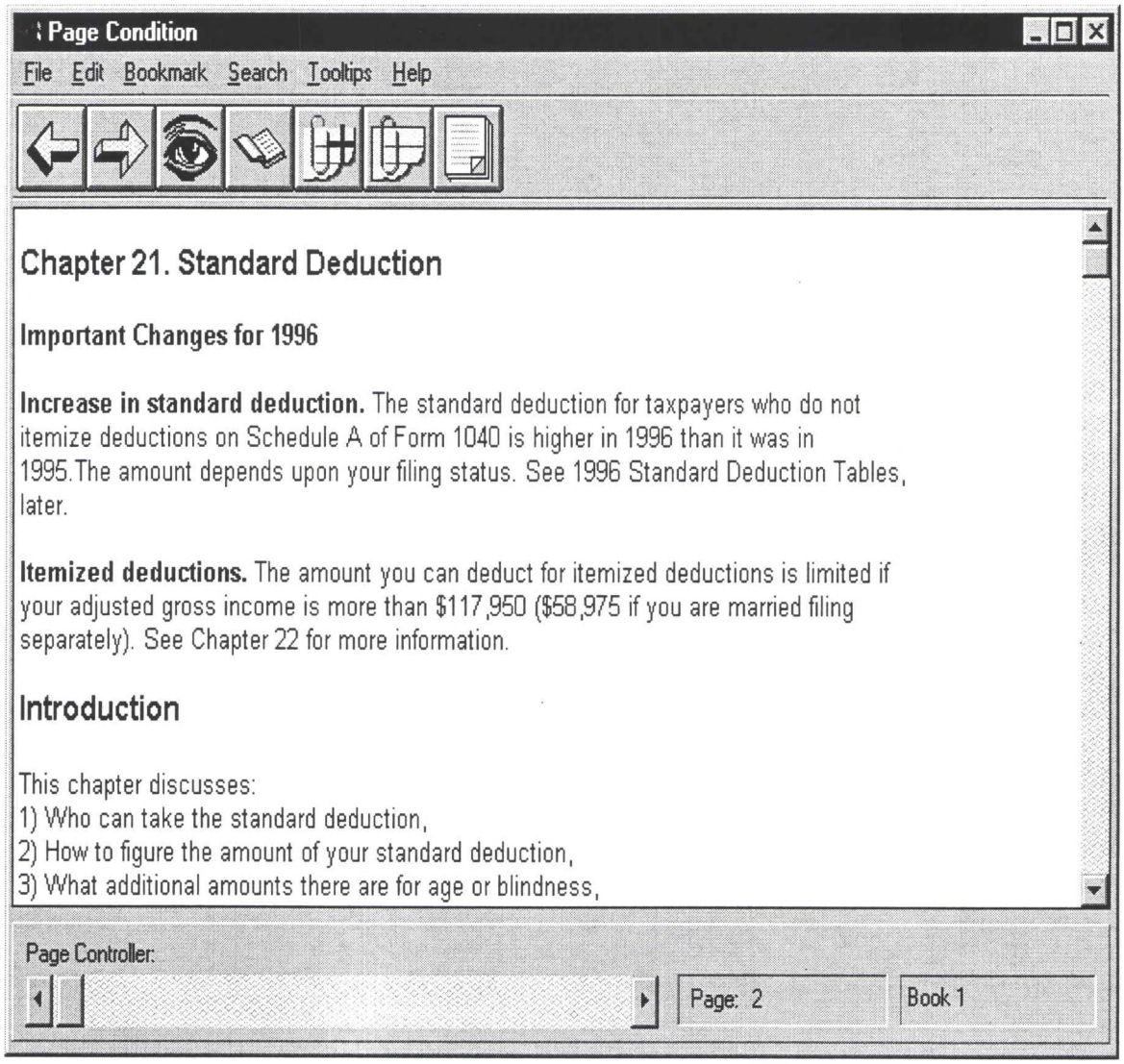


Figure 4. Sample of page as presented in the Electronic Page Flipping Mode

coordinated with the page number indicator and the text such that movement of the scrollbar changed the specific page number and text which was being viewed.

To implement the PF model, it was necessary to store the book as separate files, one file for each page. For the Electronic Page Flipping Short Page (PF-SP) version, there were 365 files that were stored and encoded. For the LP condition, there was half as many files. In the SP condition, when the document was called into the text window, all 23 lines of that page appeared on the screen in front of the subject. Therefore, it was unnecessary to utilize the vertical scrollbar to view the text (See Figure 4). However, in the Electronic Page Flipping Long Page (PF-LP) condition, the entire page did not fit in text window in front of the subject. Therefore it was necessary for the subject to move the vertical scrollbar to view the entire 46 lines of the page.

Search Tasks

The 15 search tasks were generated from a large pool of search tasks composed by the present researcher. Two judges then classified each of the search tasks according to their subjective impression along three dimensions: (1) the difficulty (low, medium, and high) of finding the correct answer using the table of contents; (2) the difficulty of finding the answer using the index; and (3) the difficulty of locating the correct answer within the text once the search engine placed the subject at the appropriate location. If there was disagreement as to the difficulty on any of one of the three factors, the judges discussed the difference and arrived at a consensus opinion.

Based upon this rating schema, fifteen tasks (Appendix A) were selected which represented a range of difficulty and which afforded the opportunity to create three

different tasks sets of equivalent difficulty. Based upon the difficulty analysis, the tasks were divided into three equally difficult sets of 5 tasks each.

Questionnaires

Two different questionnaires were constructed -- an end-of-session questionnaire and a preference questionnaire. The end-of-session questionnaire (See Appendices B-D) involved rating the particular presentation mode along the following dimensions—overall reaction to the presentation mode, finding information using the presentation mode, navigating with the presentation mode, etc. These dimensions were measured along a five-point scale.

The second questionnaire was a preference questionnaire (See Appendix E) designed to determine the subject's preferences for the different presentation modes after having participating under all conditions. The questions involved ranking the three presentation modes in terms of their overall preference. The remainder of the questions involved ranking of the three presentation modes related to factors which could be potentially responsible for the performance differences.

Apparatus

There were two computers utilized for the present study. The first computer, a Pentium 166MHz computer with a 17-inch monitor, was used to display the book in the electronic conditions. The subject's eye position was approximately 50 centimeters from the computer screen. This computer was not utilized in the HC version. The second computer was used by the experimenter to capture the times and record the correctness of each task. This computer was located to the side and slightly behind the subject.

A Pascal program was written to capture the time-based measures. This program involved the use of six keys. Keys one and two were used to start and stop the timing in the overall task timer. A third key was used to measure the time spent in formulating the search – time not spent utilizing the search tools or reading the document. This key accumulated time spent in the Formulate stage (FTime) as long as the experimenter held the key down. A fourth and fifth keys were used capture the time spent during Stage 1 and Stage 2, respectively. The fourth and fifth keys functioned in much the same way as the third key in that the time was accumulated in this stage as long as the experimenter held the key down. The sixth key was used to temporarily pause the task timing while the subject gave the answer to the query. If the answer was incorrect, the experimenter released the sixth key to initiate the start timer again. If the answer was correct, the experimenter then depressed the second key which stopped the timer allowing the Pascal program to record the time accumulated in the three stages as well as the overall time and then reset the start timer to zero in order to begin a new task.

Procedure

Subjects were run individually. Each subject participated under each presentation mode (the within-subjects factor) over three sessions with the order of the presentation modes being counterbalanced by means of a 3 x 3 Latin square. Over the three sessions, the subjects participated in a total of 15 search tasks, five tasks per session. The search tasks were divided into three equivalent sets of five tasks each (based upon the aforementioned difficulty analysis). The three task sets were balanced both across session and across document condition using a 3 x 3 Greco-Latin square procedure. The use of

this procedure resulted in each tasks set occurring equally often across sessions and across document conditions.

The first part of Session I was devoted to collection of preliminary information (i.e., informed consent (Appendix F), user background and experience survey (Appendix G) as well as the general instructions for the experiment (Appendix H). The remainder of Session I and the other two sessions were identical (Appendices I-K). First, subjects were given minimal instruction on how to use specified presentation mode. For example, in the case of PF mode, the subjects were shown the functionality of the top menu and tool bars, table of contents and index tools, shown how to manipulate the horizontal scrollbar, and how to use the forward and backward buttons. Then subjects were given 10 minutes to explore the basic features of the mode as well as the document itself. The subjects were allowed to explore any features and read any material they wished.

Upon completion of the exploration period, the subjects next performed a series of five search tasks using the specified presentation mode with either the SP or LP book. The subjects were required to go through the tasks in the order presented. They were requested to perform the tasks until they found the answer or were told to stop.

The subjects were given a maximum of 10 minutes to complete each of the five tasks. To aid in following each task, the subjects were asked to signal the start of a task by announcing out loud "Beginning Task Number X". At this time, the experimenter started the timer. The subjects were asked to clearly announce when they had found the answer and physically point to the correct answer on the computer screen. At this point the timer was stopped. Then the subjects were asked to verbalize the answer. If the answer was judged by the experimenter to be correct, the task was terminated and the time was

recorded. If the subject gave an incorrect answer, but had not exceeded the time limit, the timer was restarted and the subject was asked to continue searching for the correct answer. Each task continued until, in the experimenter's judgment, the subject found the information or the time limit was reached. If the subjects failed to complete the task within the required time limit, the task was recorded as a failure and the subjects were assigned the maximum time.

Immediately following the last task of each session, subjects completed the end-of-session questionnaire indicating their subjective impression of the presentation mode. Lastly, the evaluator conducted an open-ended, subjective debriefing of the subject regarding their general impressions.

The above procedure was repeated for each session. In addition, at the conclusion of the final part of Session 3, subjects completed the preference questionnaire which involved a comparative evaluation of the three presentation modes and were then given a debriefing sheet (Appendix L).

Dependent Variables

The primary dependent variables were the percentage of tasks completed successfully and the average time to complete the task when correct. In addition, the overall time-based measure (TskTime) was decomposed into three separate components. These components included the total time the subject spent formulating the task (FTime), utilizing the search tools (S1Time) , and searching the document for the correct answer (S2Time).

CHAPTER 3

RESULTS

The presentation of the results is organized into five major sections pertaining to the type of data which was collected. The first section characterizes the subjects and point out differences related to the previous Biers et al. usability study (1995). The second section presents the performance data which incorporates traditional measures of accuracy and completion time. The third section examines the end-of-session subjective data questionnaires that included questions about difficulty and frustration of the presentation modes as well as salient dimensions in which the modes were expected to differ. The fourth section considers the results of the end-of-study preference questionnaire in which the subjects indicated their subjective preference for the three presentation modes. The final section presents the subject's responses to questions dealing with potential confounding differences between the computer and hardcopy modes (i.e. manipulation checks).

All results were analyzed separately for short (SP) and long pages (LP). The presentation modes were structured so that the subject was sent to the top of the page that contained the index term on which they searched. In the SP condition, the index term always appeared on the screen although the answer might not appear on that page. However for LP, although the subjects were also sent to the top of the page containing the

index term, it was a different place than for SP. In addition, the index term might have been hidden from the subject in the Electronic Page Flipping (PF) and Electronic Scrolling (ES) condition, but not in the hardcopy (HC). Therefore the decision was made to analyze the data separately for SP and LP. By doing so, it decreased the power of to detect significant differences and did not permit a direct test of the interaction.[1]

Analysis of the Subject Profile Questionnaire

Table 5 summarizes the subject profile information for subjects under the SP and LP condition. Since most of the questions represented ordinal data, the median response was used to characterize the subjects. Inspection of Table 3 reveals that most of the subjects were college graduates, had 6-7 years of PC Usage, used the PC a few times a week at home and at work, and had used MS-Windows for 3-4 years. However, there were some noticeable differences between the subjects in the two page length conditions. Subjects in the LP condition used a PC a slightly more at work, were more comfortable using MS-Windows, were more positive about reading text on the computer screen, and were more likely to have read tax information online or in its HC format than subjects in the SP condition.

[1] The analyses were also conducted utilizing a 2 (page length) x 3 (presentation mode) Analysis of Variance, but the statistical conclusions did not change even with greater power.

Table 5

Median Response on Selected Demographics Questionnaire

Questions	Page Length	
	Short Pages	Long Pages
Education Level	College Graduate	College Graduate
Length of PC Usage	6-7years	6-7years
Frequency of PC Usage at Home	Few times a week	Few times a week
Frequency of PC Usage at Work	Few times a week	Between a few times a week and once a day
Frequency of MS-Windows Usage	3-4 Years	3-4 Years
Comfort with MS Windows	Somewhat	Very
Like Reading text information on screen	Neutral	Like
Like Reading Help Feature	Neutral	Neutral
Read tax information online or hardcopy	Never	Less than once a month

Based upon a classification schema utilized by Biers et al. (1995), an attempt was then made to classify the subjects into broad experience categories. In the Biers (1995) et al. study, there were three variables that appeared to differentiate the subject's level of experience. The first variable was the maximum frequency of PC usage at home or at work. The dividing line in the Biers et al. (1995) study seemed to be between those who used the PC once or more a day versus those who used the PC less than once a day. Using the same criteria (See Table 6), the majority of subjects under LP spent once a day or more using the computer at home or at work. However for SP, the subjects were evenly split between the two usage categories. Thus the level of computer usage was much higher for subjects assigned the LP condition.

Table 6

Reported Frequency of Maximum Computer Use - Home or Work

Frequency of Use	Number of Subjects	
	Short Pages	Long Pages
Less than Once a Day	8	4
Once a Day or More	7	11

The second variable that seemed to differentiate the experience level in the Biers et al. (1995) study was the number of advanced Windows features that were utilized. In the user profile of the present study, these advanced features included using the file manager or windows explorer, running multiple windows applications, moving or resizing the windows, copying and pasting from one application to another, and right-clicking with the mouse. The number of advanced features utilized was divided into three levels: 0-1 features, 2-3 features, and 4-5 features (See Table 7). In the Biers et al. study (1995), no one used more than 2-3 features. However, inspection of Table 5 reveals that over 50% percent of the subjects (n=17) used 4 to 5 features indicating the subjects in the present study were more experienced than in the Biers et al. (1995) usability test. Secondly, the majority of LP subjects (11 of 15) in this study used 4-5 features whereas with SP, the subjects were evenly split among the three levels. This analysis was consistent with the frequency of PC usage indicating that the subjects in the LP condition were more experienced than the SP condition.

Table 7

Reported Number of Windows Feature Used

Windows Feature Use	Number of Subjects	
	Short Pages	Long Pages
0-1 Features	4	0
2-3 Features	5	4
4-5 Features	6	11

The third variable that seemed to differentiate the experience level of subjects in the Biers et al. (1995) usability study was the subject's comfort level with using MS-Windows. The subject's comfort level was divided into three categories: Not Very Comfortable, Somewhat Comfortable, and Very Comfortable (See Table 8). In the LP condition, the majority of subjects (13 of 15) felt very comfortable with MS-Windows. But in contrast, the majority of subjects in the SP (9 of 15) condition felt at best somewhat comfortable with MS-Windows. Thus, as with the previous two variables, there were experience differences between the two page length conditions.

In comparison to the Biers et al. (1995) study, the subjects in the present study were much more comfortable using MS-Windows. In the prior study, the majority of subjects were somewhat comfortable whereas in the present study, two-thirds of the subjects were very comfortable using Windows.

Table 8

Reported Comfort with MS-Windows

Comfort Level	Number of Subjects	
	Short Pages	Long Pages
Not Very	1	1
Somewhat	8	1
Very	6	13

Based upon these three aforementioned variables, the subjects were classified into one of three categories – beginner, intermediate, and advanced using the same criteria of the Biers et al. (1995) study (See Table 9). Consistent with the Biers et al. (1995) study, a beginner was defined as using 0 to 1 advanced features of MS-Windows, having used a computer less than once a day, and not very/somewhat comfortable with MS-Windows. An advanced subject was classified as one who used 4-5 advanced features of Windows, having used a computer one or more times a day, and felt comfortable using MS-Windows. The other subjects were classified as intermediates.

Table 9

Number of Subjects classified on Computer Experience

Experience Level	Number of Subjects	
	Short Pages	Long Pages
Beginner	4	1
Intermediate	8	4
Advanced	3	10

From this reclassification, it can be seen that the subjects assigned to the two page lengths conditions were not equivalent despite using random assignment (χ^2=6.90, p=.032). The subjects assigned to the LP condition were significantly much more

advanced in their knowledge and experience with MS-Windows than the subjects in the SP condition.

There are striking differences between the present study and the Biers et al. (1995) usability test. First, in the Biers et al. (1995) usability test, there were only beginner and intermediate experienced subjects whereas in the present study almost 50% of the subjects were considered advanced subjects. Secondly, 70% of the subjects in the Biers et al. (1995) study were classified as beginner subjects whereas in the present study only 17% were considered beginners. Thus the subjects of the present study were much more experienced.

Analysis of Performance

The performance data was analyzed at two different levels. The first level of analysis consisted of aggregating the data across all individual tasks to yield overall performance measures for each mode condition. The second level of analysis examined performance on an individual task-by-task basis.

Overall Performance

The data was collapsed across individual tasks to yield performance measures based on the five tasks per presentation mode condition. There were five dependent measures -- Percent Correct and four time-based measures. Percent Correct measured the percentage of tasks completed out of five tasks. The successful completion of a task was defined as the subject finding the correct answer within a time limit of 10 minutes.

There were four time-based measures derived from the tasks in which a correct answer was given. First, overall Task Time (TskTime) was measured from the initiation of the task until the subject found the correct answer. In addition, overall task time was

broken into three components. The first component of overall task time was Formulate time (FTime) which was measured when the subject read the task statement and formulated how he/she would search for a task. The second component was Stage 1 time (S1Time) which measured the time the subject utilized the search tools to find the answer. The subject had two search tools at his/her disposal in order to aid in finding the answer-- the table of contents and/or the index. The third component was Stage 2 time (S2Time) which is the time the subject spent browsing the document for the answer. If an incorrect answer was given, the experimenter instructed the subject to keep searching for the correct answer. In this case the subject had a number options available. The subject could revisit any of the three stages (FTime, S1Time, and S2Time) until the correct answer was found or time had expired. The number of times each stage was revisited was not recorded.

Table 10 presents the raw means for the six conditions formed from the factorial combination of presentation mode and page length on each of the five dependent measures. Inspection of Table 10 reveals that the standard deviations for the time-based measures were very large relative to the difference in means which suggested large error variability. Second, not shown in this table, the distributions of these time-based measures were highly positively skewed. In all cases, the Shapiro-Wilkes (a test for skewness) was significant (all $p's<.01$). This suggests the option of either transforming the data or performing a non-parametric statistical test.

Table 10

Means and Standard Deviations for percent correct and the four time-based measures

		Page Length					
		SP			LP		
Measure		HC	ES	PF	HC	ES	PF
PctCor	Mean	97.33	94.67	92.00	96.00	98.67	89.33
	Std Dev	7.04	9.15	12.65	8.28	5.16	12.80
F1Time	Mean	14.05	15.98	18.28	13.77	16.41	13.55
	Std Dev	4.67	4.78	9.43	3.56	7.92	5.20
S1Time	Mean	35.99	32.79	32.82	27.87	29.33	23.44
	Std Dev	10.30	10.77	14.82	12.81	13.68	11.90
S2Time	Mean	94.87	100.85	74.32	78.17	87.97	88.95
	Std Dev	45.41	40.16	34.49	37.09	31.08	44.00
TskTime	Mean	144.90	149.62	125.42	119.81	133.71	125.94
	Std Dev	56.63	50.02	51.04	47.52	43.50	51.40

As a consequence of the above, the data was analyzed in several different ways- Analysis of Variance (ANOVA) on the raw non-transformed data, ANOVA on the transformed data, non-parametric analysis using the Friedman analysis of ranks, and Analysis of Covariance (ANCOVA) to reduce error variability. There were three attributes that served as covariates -- task set performance, subjective difficulty, and subjective frustration.

In transforming the data, the transformations recommended by the SPSS Examine procedure were employed. In somewhat of an unusual situation, a different

transformation was recommended for each time-based measure. Table 11 shows the transformations that were utilized.

Table 11

Transformations Recommended by SPSS

Dependent Measure	Transformation Used
FTime	Reciprocal
S1Time	Log
S2Time	1/Square Root
TskTime	Square Root

In conducting the ANCOVAs, it was critical to select variables that change across time since the independent variable was manipulated within-subjects. A constant covariate, such as prior user experience, would have no impact on the effect of presentation mode. In this data set, task set performance, subjective difficulty, and subjective frustration changed as the presentation mode condition varied.

In addition, task set performance, subjective difficulty, and subjective frustration were selected as covariates there was some suggestion that performance varied as a function of these variables. First, an ANOVA was performed on the time-based measures to check whether performance differed as a function of Task Set (1,2,3). Given the procedure to assign tasks to task sets, there should be no reason to expect a difference. For S1Time, there was a significant difference in the three task sets (F=5.663, p<0.004). In addition, the effect of task set approached significance for overall task set (F=2.579 , p<0.077). Therefore, to control for this potentially confounding variable, average task set performance was covaried.

The measures of subjective difficulty and frustration were correlated with the time-based measures. It was found that the average difficulty significantly correlated with S2Time (difficulty: r=-0.137, p=0.046); average difficulty and frustration were correlated with S1Time (difficulty: r=-0.189, p=0.006; frustration: r=-0.201, p=0.003) and overall task time (difficulty: r=-0.166 , p=0.015; frustration: r=-0.147, p=0.032). In ANCOVA, measures correlated with the dependent variable reduced error variance and increase the sensitivity of the analysis. Therefore, subjective difficulty and frustration were also treated as covariates.

Table 12 summarizes the results these analysis. In this table, the analysis and the test used are displayed across the top and the dependent measures are listed on the side. The body of the table contains the probability values for the effect of presentation mode for the appropriate statistical procedures. When the presentation mode effect was significant, the table shows the locus of significance (i.e. the pairwise comparisons that were significant). To facilitate locating significant effects in the table, the cells in which significance was attained are shaded in gray.

Table 12

Results of Analysis of the Performance Measures using ANOVA and ANCOVA

Short Pages (p values)

D.V.	Raw ANOVA	Transformed ANOVA	Non-Parametric Friedman	Task Set ANCOVA	Difficulty ANCOVA	Frustration ANCOVA
PctCorr	0.381	N/A	0.446	N/A	N/A	N/A
FTime	0.123	0.192	0.041 HC<ES HC=PF, ES=PF	0.068	0.098	0.095
S1Time	0.704	0.238	0.420	0.673	0.611	0.714
S2Time	0.146	0.335	0.085	0.143	0.233	0.162
TskTime	0.375	0.414	0.344	0.317	0.490	0.394

Long Pages (p values)

D.V.	Raw ANOVA	Transformed ANOVA	Non-Parametric Friedman	Task Set ANCOVA	Difficulty ANCOVA	Frustration ANCOVA
PctCorr	0.040 ES>PF HC=ES, HC=PF	N/A	0.025 ES>PF HC=ES, HC=PF	N/A	N/A	N/A
FTime	0.112	0.123	0.449	0.109	0.139	0.177
S1Time	0.428	0.357	0.085	0.321	0.473	0.510
S2Time	0.683	0.946	0.549	0.681	0.706	0.711
TskTime	0.719	0.525	0.819	0.641	0.773	0.798

Note: Numbers indicate obtained probability of statistic due to chance.

Inspection of Table 12 shows that the effect of presentation mode was non-significant in the majority of cases, irrespective of the analysis approach. Given the variety of analysis approaches, this table clearly documents that a mode effect was not evident in the present data, at least at the level of the aggregated data. For purposes of presentation of the remainder of the results, the decision was made to focus upon the ANOVA's performed on the untransformed data and the non-parametric statistics. This decision was made for two reasons. First, since the normality assumption was violated, a non-parametric analysis was most appropriate. Secondly, presentation of the raw untransformed data gives the reader a clearer picture of the results using the original metrics. In any case, choosing these two approaches makes no difference in the statistical conclusion except in one case.

Given that the pattern of the means was sometimes different using the raw and ranked data, the graphs depicting the performance results are divided into two panels: the upper panel presents the means for the raw data and the lower panel depicts the means for the ranked data. Each aggregated performance measure is now described below.

Percent Correct

Inspection of Figure 5 revealed that the percent correct ranges from 92% to 97% in the SP condition and from 89% to 99% for the LP condition. For both SP and LP, the PF mode resulted in the lowest percent correct. Analysis of both the raw and ranked data revealed that the differences were not significant for SP; however for LP, these differences were significant. The two analysis (see Table 9) revealed the PF (M=89.33%) mode had a significantly lower percent correct than did the ES (M=98.67%) mode for LP.

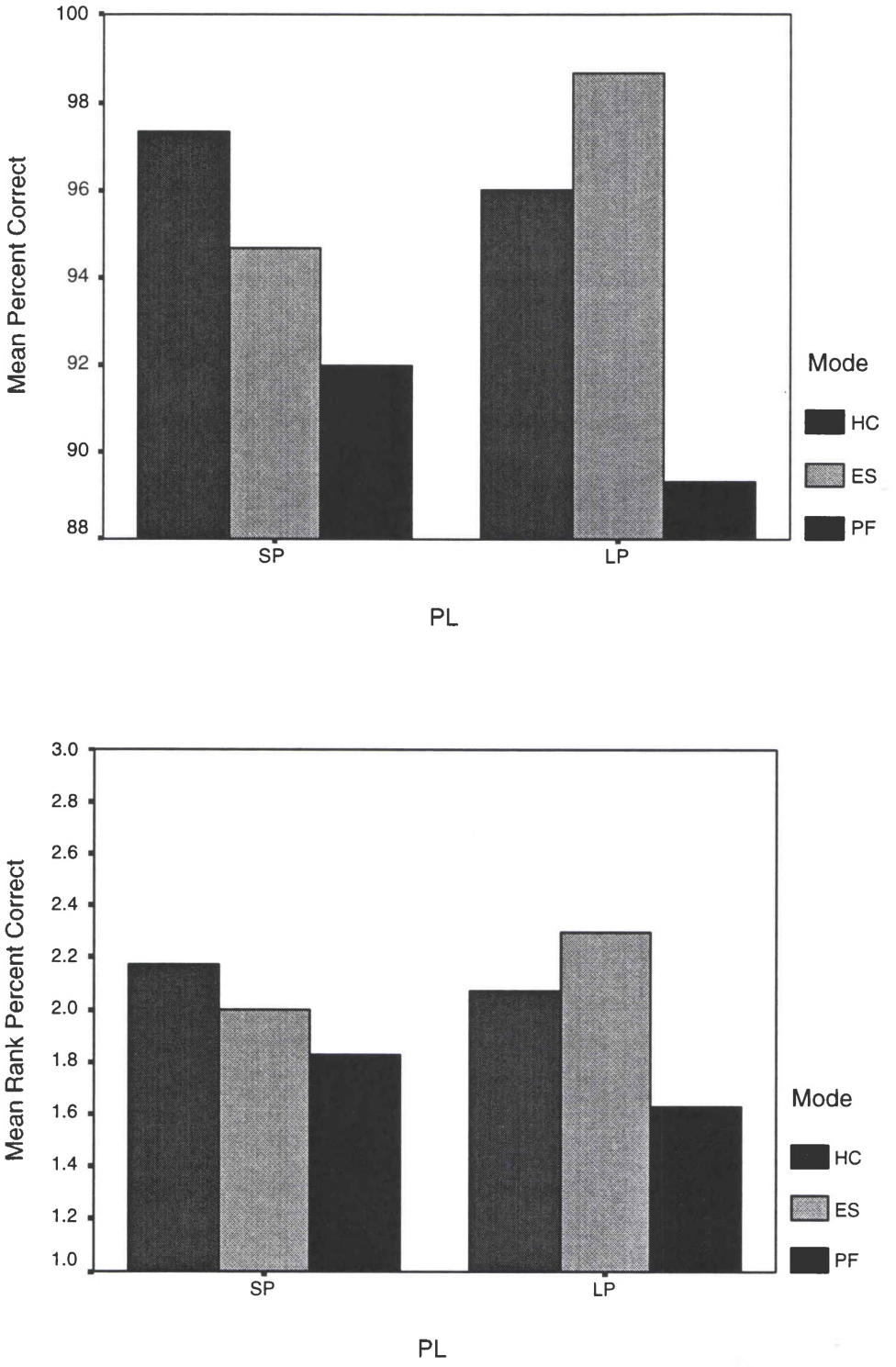


Figure 5. Raw (Top) and Ranked (Bottom) Means for Percent Correct

Although, no specific predictions were advanced with regard to percent correct, the results for SP were consistent with what one might expect in the amount of text displayed hypothesis. Namely with the amount of text displayed hypothesis held constant there should be no significant differences among the three modes. The results with regard to LP was more consistent with the book metaphor hypothesis where the PF mode should result in the poorest performance. However, contrary to the same hypothesis, HC did not result in a superior performance to ES. However, the failure of the HC to result in superior performance under LP may be due to the ceiling effect.

Formulate Time

Formulate Time (FTime) represented the time to formulate the task. One would not expect that the mode or page length would effect the time that the subject spent formulating the problem. Figure 6 plots the FTime for both the raw non-transformed data and the ranked data. Inspection of Figure 6 showed that the means varied from 13 to 18 seconds where the longest time spent in Formulate Stage being with the SP-PF condition. Note that the ranked data showed a different pattern because of the effect of skewness was removed from the data.

Inspection of Table 12 reveals that there were no significant differences in FTime for all analysis approaches with the exception of the non-parametric analysis for SP. Analyses of simple comparisons on the ranked data showed subjects spent significantly more time formulating the search strategy under both the ES and PF modes than the HC mode. There was no a priori reason to expect any difference in this measure.

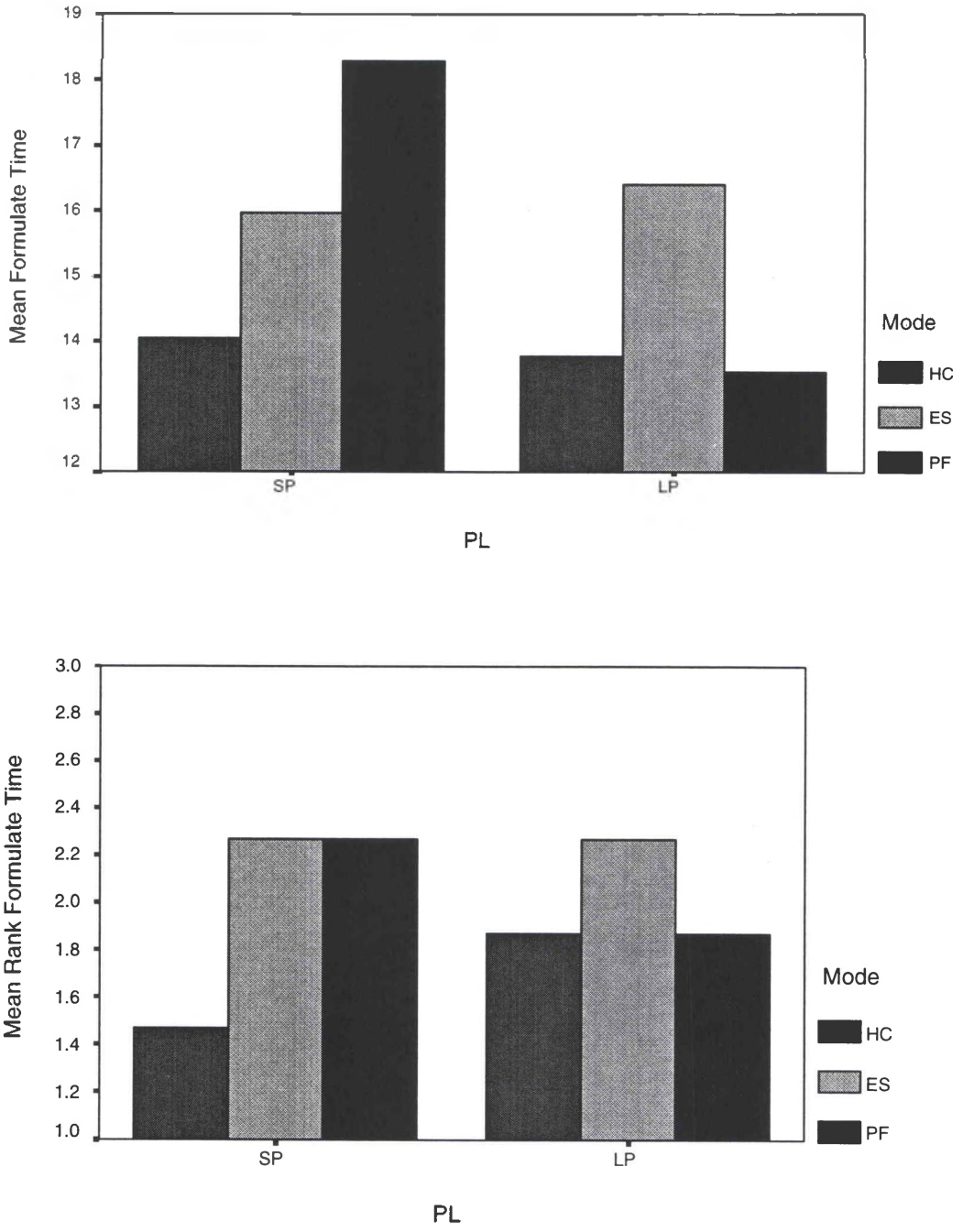


Figure 6. Raw (Top) and Ranked (Bottom) Means for Formulate Time

Stage 1 Time

The mode conditions were not expected to show any difference in Stage 1 time (S1Time) since a low level interface search tool was created in which the index and table of contents were identical across the three presentation modes. In addition, the amount of text displayed in the index at any one time was held constant in all three conditions.

Analysis of the data revealed that for both SP and LP across all analysis approaches, there was no significant differences in S1Time. Figure 7 shows that the mean times ranged from 24 seconds to 36 seconds across all six conditions. A closer look at Figure 7 reveals that it took longer to search the index and table of contents for SP than LP but there was no reason to expect this difference. The mean S1 time for SP was approximately 34 seconds while the mean S1 Time for LP was approximately 27 seconds.

Stage 2 Time

The majority of the hypotheses were formulated relative to Stage 2 Time (S2Time). S2Time referred to the amount of time that the subjects spent browsing through the text itself after using a search tool. It was expected that the effect of presentation mode would be different for SP and LP, but the exact nature of the interaction would be different for the amount of text displayed and the book metaphor hypothesis. Figure 8 presents the means for each of the six different conditions. For SP, the means ranged from 72 seconds to 100 seconds; for LP, the means ranged from 78 to 90 seconds. However, despite these apparent differences, the effect was not significant for any analysis approach for either SP or LP. The failure to attain significant mode effect for either SP or LP is totally inconsistent with the book metaphor hypothesis. With regard to

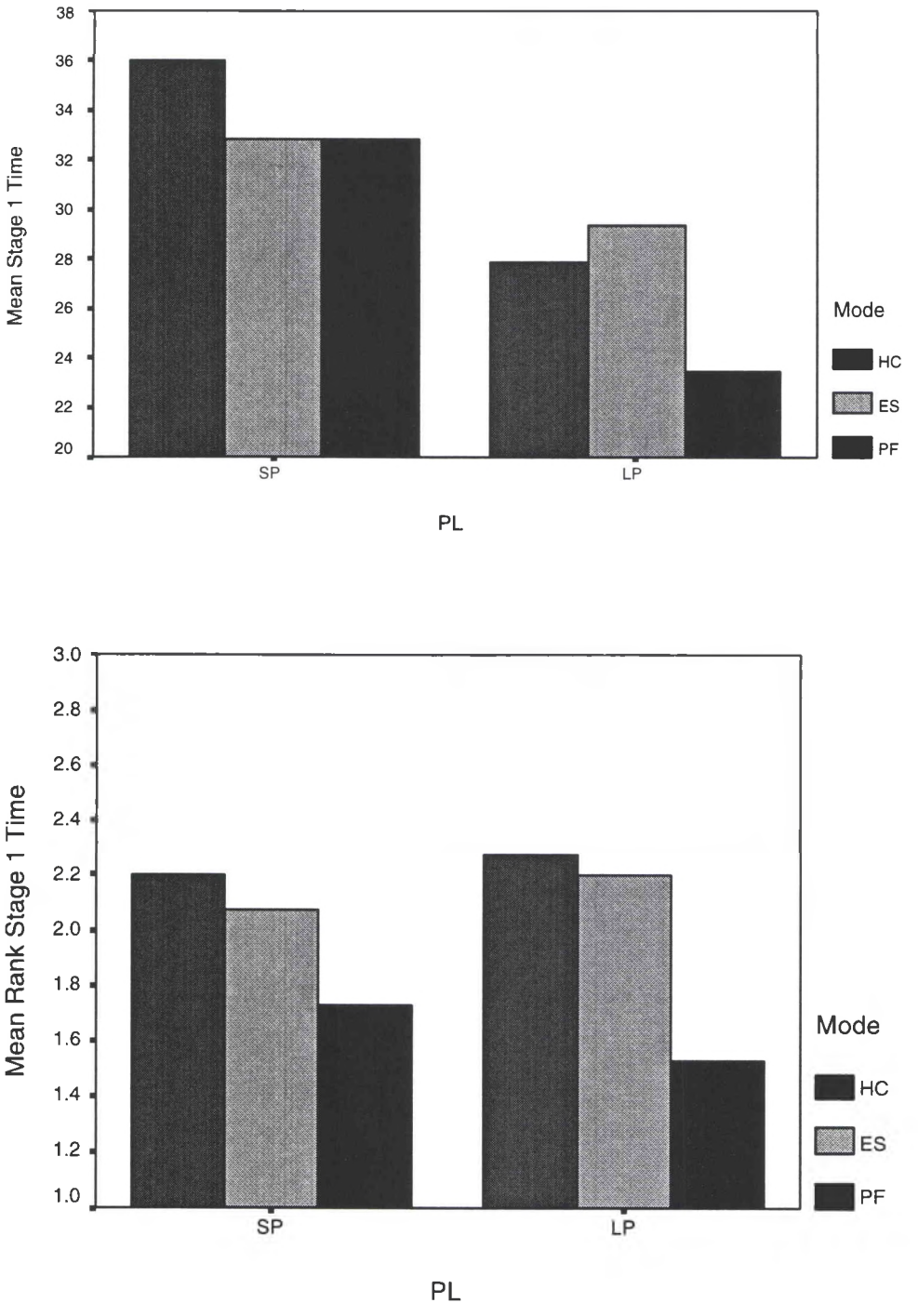


Figure 7. Raw (Top) and Ranked (Bottom) Means for Stage 1 Time

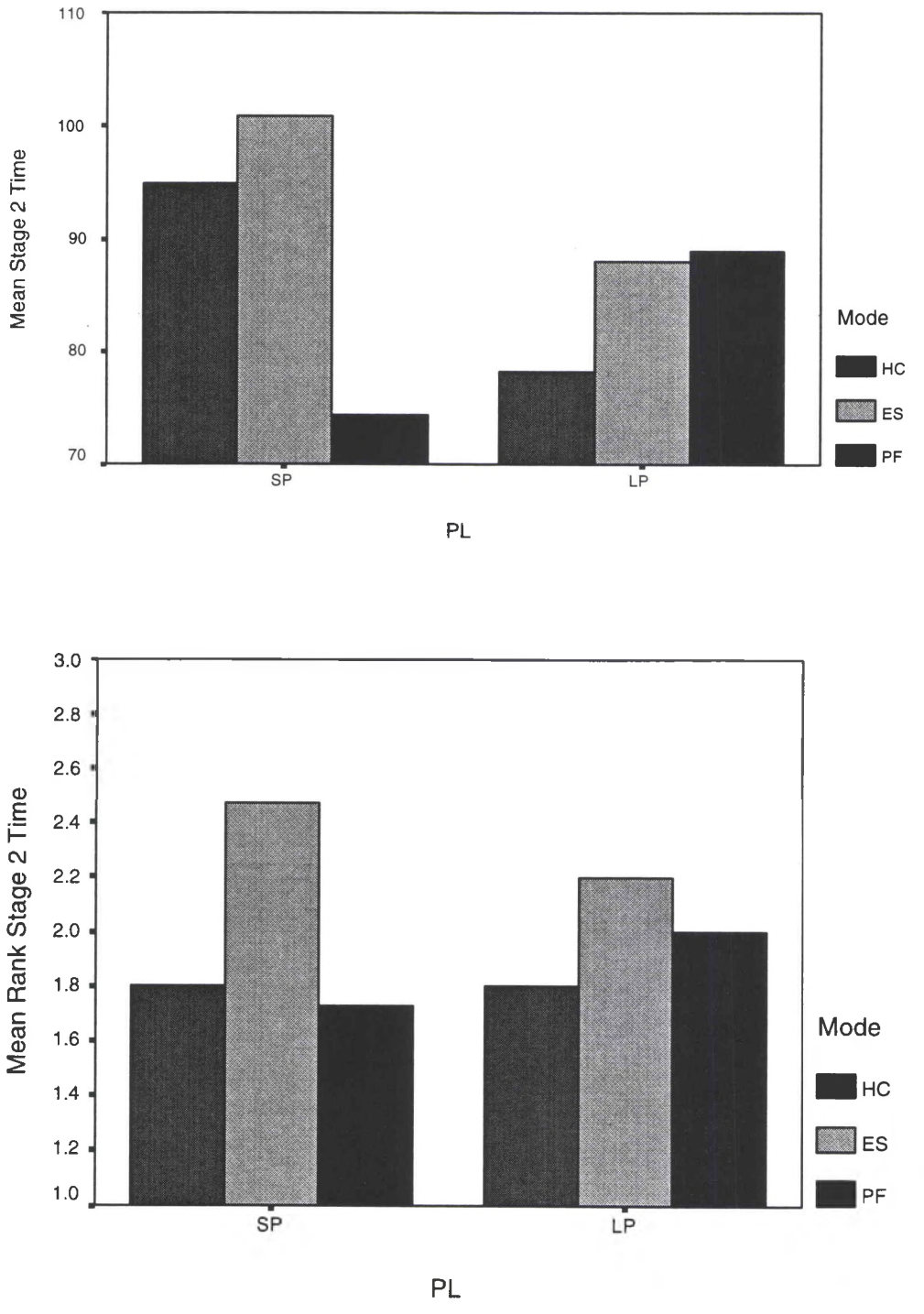


Figure 8. Raw (Top) and Ranked (Bottom) Means for Stage 2 Time

the amount of text displayed hypothesis, the lack of significance for mode was expected for SP but not for LP.

Overall Task Time

The dependent measure, overall task time (TskTime), was the sum of FTime, S1Time, and S2Time. Since overall task time included S2Time, the predictions made for S2Time were expected to also hold for overall task time. As shown in Figure 9, the mean overall task times ranged from 120 to 150 seconds. However, as with S2Time, none of these differences were significant.

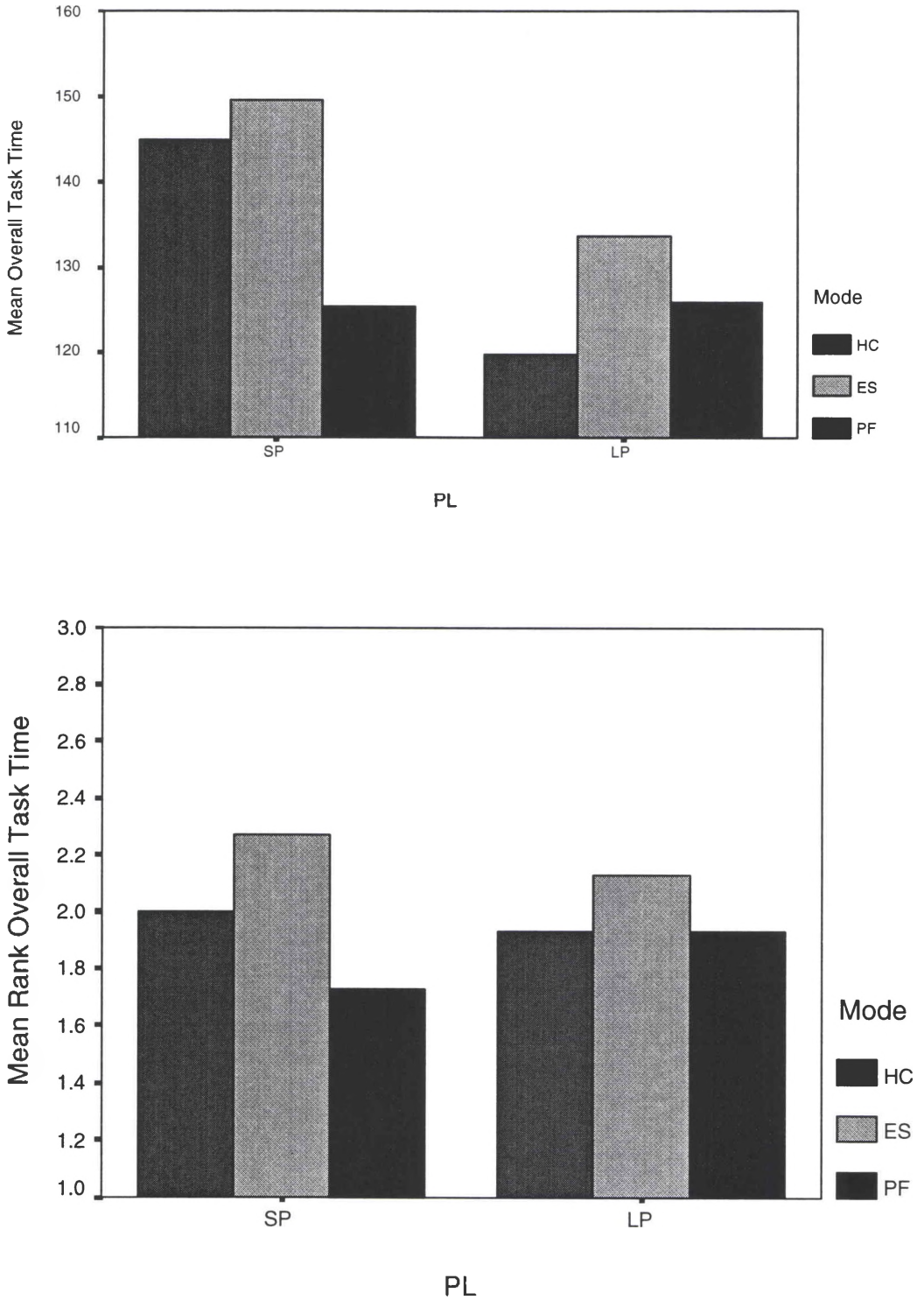


Figure 9. Raw (Top) and Ranked (Bottom) Means for Overall Task Time

Task by Task Analysis

The overall performance data could have become obscured by the fact that the data was aggregated across individual tasks. Perhaps there were mode differences associated with particular tasks and not others. Therefore each of the fifteen tasks was analyzed separately using both the non-transformed raw data (ANOVA) and the ranked data (Kruskal-Wallis). In the analysis of the data on a task-by-task basis, the document mode factor became a between-subject variable and this resulted in a loss of power. Appendix A contains a detailed description of the individual search tasks.

Table 13 summarizes the significance of the results for all dependent measures for each of the fifteen tasks. As shown in Table 13, the majority of the analyses showed no significant difference as a function of presentation mode for either SP or LP. When there was a significant difference (only 8 of the 15 tasks), the effect of mode was not consistent across the dependent measures. The following is a discussion of the tasks in which a significant difference was attained.

Task 1 : Real Estate Taxes

There was an effect of presentation mode on both S2Time and overall task time for LP and not SP. For LP, the subjects took significantly longer using the PF (S2Time M=63.00;TskTime M=102.25) mode than either the ES (S2Time M=32.00; TskTime M=66.60) or the HC (S2Time M=44.40; TskTime M=72.60) modes for S2Time (raw data only) and overall task time (both raw and ranked data). However, using the ranked data for S2Time, the results were different-- the subjects took significantly longer with the ES mode than the HC mode whereas with the raw data, they were equivalent.

Table 13

Results of Task-by-Task Analysis using ANOVAs

Task	Analy	Percent Correct Raw	Percent Correct Rank	Ftime Raw	Ftime Rank	S1Time Raw	S1Time Rank	S2 Time Raw	S2 Time Rank	TskTime Raw	TskTime Rank
1	SP	1.000	1.000	0.915	0.915	0.633	0.722	0.136	0.203	0.269	0.779
	LP	0.397	0.368	0.109	0.109	0.175	0.219	0.016 HC=ES<PF	0.044 HC<ES (.056), HC=PF, ES=PF	0.006 HC=ES<PF	0.036 HC=ES<PF
2	SP	1.000	1.000	0.612	0.862	0.016 PF<HC, HC=ES, ES=PF	0.059 PF<HC(.056), HC=ES, ES=PF	0.226	0.230	0.449	0.482
	LP	1.000	1.000	0.705	0.675	0.782	0.934	0.296	0.281	0.390	0.612
3	SP	0.397	0.368	0.424	0.233	0.023 ES<HC, HC=PF, ES=PF	0.044 ES<HC, HC=PF, ES=PF	0.458	0.610	0.363	0.309
	LP	1.000	1.000	0.705	0.675	0.782	0.934	0.296	0.281	0.390	0.612
4	SP	1.000	1.000	0.804	0.684	0.758	0.778	0.787	0.763	0.788	0.619
	LP	1.000	1.000	0.975	0.719	0.007 PF<HC, HC=ES , ES=PF	0.024 PF<HC, HC=ES, ES=PF	0.856	0.932	0.900	0.853
5	SP	0.619	0.584	0.391	0.457	0.339	0.509	0.213	0.112	0.271	0.358
	LP	0.335	0.335	0.331	0.211	0.440	0.730	0.034 HC<PF, HC=ES, ES=PF	0.029 HC<ES, HC=PF, ES=PF	0.120	0.090
6	SP	1.000	1.000	0.487	0.431	0.134	0.144	0.225	0.294	0.202	0.23
	LP	1.000	1.000	0.677	0.408	0.559	0.615	0.543	0.543	0.507	0.394
7	SP	0.397	0.368	0.235	0.241	0.420	0.682	0.462	0.617	0.426	0.476
	LP	1.000	1.000	0.587	0.618	0.427	0.334	0.658	0.584	0.796	0.694

Table 13 (continued)

Task	Percent Correct		Ftime		S1Time		S2 Time		TskTime	
	Raw	Rank	Raw	Rank	Raw	Rank	Raw	Rank	Raw	Rank
8 SP	1.000	1.000	0.542	0.787	0.372	0.479	0.208	0.048 PF<ES, HC=ES, HC=PF	0.238	0.156
LP	1.000	1.000	0.227	0.375	0.587	0.746	0.022 HC<PF, HC=ES, ES=PF	0.076	0.080	0.179
9 SP	1.000	1.000	0.924	0.948	0.519	0.430	0.474	0.326	0.458	0.33
LP	1.000	1.000	0.791	0.704	0.632	0.468	0.309	0.566	0.384	0.827
10 SP	1.000	1.000	0.422	0.551	0.230	0.129	0.289	0.307	0.268	0.206
LP	1.000	1.000	0.972	0.897	0.567	0.783	0.468	0.887	0.487	0.914
11 SP	0.110	0.116	0.547	0.677	0.474	0.869	0.347	0.267	0.448	0.339
LP	0.397	0.368	0.259	0.168	0.293	0.362	0.748	0.764	0.439	0.617
12 SP	0.110	0.116	0.891	0.996	0.402	0.540	0.490	0.313	0.520	0.566
LP	0.397	0.368	0.194	0.063	0.303	0.418	0.017 HC=PF<ES	0.131	0.011 HC=PF<ES	0.112
13 SP	1.000	1.000	0.696	0.851	0.751	0.883	0.146	0.067	0.210	0.528
LP	1.000	1.000	0.636	0.530	0.695	0.832	0.293	0.121	0.332	0.150
14 SP	1.000	1.000	0.273	0.337	0.065	0.060	0.291	0.364	0.167	0.137
LP	1.000	1.000	0.264	0.295	0.919	0.648	0.533	0.543	0.770	0.675
15 SP	0.100	0.108	0.397	0.632	0.166	0.133	0.584	0.601	0.420	0.355
LP	0.016 HC=ES>PF	0.03 ES>PF, HC=ES, HC=PF	0.629	0.881	0.297	0.434	0.513	0.323	0.626	0.489

Note: Greater than(>) indicates longer search times; Numbers indicate obtained probability of statistic due to chance

The failure to find any significant differences with SP is consistent with the amount of text displayed hypothesis because the amount of text information displayed to the subject was equivalent. The fact that the HC results in shorter times than PF for LP is also consistent with this hypothesis. However contrary to the amount of text displayed hypothesis, HC was not faster than ES. With LP, significance was predicted with both hypothesis, but the pattern of the results did not conform to either hypothesis.

Task 2: Charitable Organization

There was an effect of presentation mode on S1Time for SP using the ranked data. Analysis of simple comparisons revealed that subjects spent more time searching the table of contents and index with the HC (M=48.60) mode than the PF (M=11.20) mode. Though no hypothesis was advanced with S1Time, this result suggested that the subjects had difficulty using the HC search tools. There were no significant differences between the HC and ES modes and between the ES and PF modes.

Task 3: Itemized Deduction Limitation

For Task 3, the analysis of both the raw and ranked data was consistent in showing an effect of presentation mode for S1Time. In both analyses, the only significant difference was that between HC and ES - it took longer to find the appropriate search terms with HC (M=29.40) than ES (M=15.60). Once again, this suggests that the subjects had difficulty using the HC search tools.

Task 4 : Nursing Home Deduction

For Task 4, there was a significant effect of presentation mode on S1Time using both the raw and ranked data. However, unlike Task 2 and 3, the effect was for LP and not SP. As with Task 2 and 3, there appeared to be a problem with using the search tools

with HC. Analysis of simple comparisons revealed that subjects took longer to find the index terms in Stage 1 with the HC (M=27.40) mode than the PF (M=9.60) mode.

Task 5: Refinancing the Home

In Task 5, significant differences in the three presentation modes were obtained for S2Time for both raw and ranked data with LP only. The raw and ranked data were consistent in showing that subjects spent more time in Stage 2 with the PF (M=249.00) than the HC (M=105.25) mode. The results parallel those of Task 1 with the exception that there was no difference between ES and PF.

Task 8: Dependent and Medical Deduction

There were significant mode differences with S2Time in Task 8. However the two analysis approaches were inconsistent in the pattern of significance. With the raw data, there was a significant difference for LP and not SP. Using the ranked analysis, there was a significant difference with SP and not LP. For LP, analysis of the raw data revealed significantly more time spent in Stage 2 with the PF (M=63.00) mode than the HC (M=23.40) mode. For SP, analysis of the ranked data showed that subjects took longer in Stage 2 with the ES mode than with the PF mode. The inconsistency in the results for Task 8 across the two analyses approaches makes the task difficult to interpret.

Task 12: Educational Course

There was also a significant effect of presentation mode on S2Time and overall task time for LP using the raw data. For the raw data, subjects spent significantly shorter times in Stage 2 and in the total task with the HC (M=62.20) and PF (M=52.75) modes than the ES (M=189.60) mode. For Tasks 2, 5, and 12, the HC times were consistently

shorter than either ES or PF times for both S2Time and overall task time under the LP condition.

Task 15: Eye Condition

There was a significant effect of presentation mode on answer correctness for LP using both the raw and ranked data. For the raw data, subjects were more likely to be correct with the ES (M=1.00) and HC (M=1.00) modes than the PF (M=0.40) mode. There was no significant difference between the HC and ES modes. For the ranked data, the only a significant difference was between the ES and PF where the subjects were more likely to answer the questions correctly with the ES mode than the PF mode.

Analysis of the End-of-Session-Questionnaires

At the end of each session, the subjects were asked a set of parallel questions. There were 5 questions dealing with the subject's overall impression of the presentation modes, difficulty of finding specific information, difficulty in using the different search tools, difficulty with location and navigation, the value of the mode, and the difficulty and frustration of the search tasks themselves. Appendices H-J shows the end of session questionnaires in their entirety.

Table 14 shows the means and standard deviation for each presentation mode with higher numbers indicating a more positive evaluation. Table 15 summarizes the significance of the ANOVA's performed on the data.

For the majority of questions there were no differences between the presentation modes for either SP or LP. When significant differences were obtained, the electronic scrolling and the page flipping modes were more stimulating, flexible, and less difficult to use. In none of the questions did the subjects rate the hardcopy mode across page.

Table 14

Means and Standard Deviations for the End-of-Session Questionnaires

Question 1A-Overall, Terrible/Wonderful

	Short		Long	
	Mean	SD	Mean	SD
HC	3.13	.83	3.33	.98
ES	3.67	.90	3.93	.26
PF	3.73	.96	3.67	.98

Question 1B – Overall, Frustrating/Satisfying

	Short		Long	
	Mean	SD	Mean	SD
HC	3.20	1.08	3.53	1.06
ES	3.60	1.18	3.73	0.46
PF	3.67	1.11	3.60	1.06

Question 1C- Overall, Dull/Stimulating

	Short		Long	
	Mean	SD	Mean	SD
HC	2.33	0.81	2.47	1.06
ES	3.13	0.74	3.40	0.83
PF	3.33	0.82	3.20	1.37

Question 1D - Overall, Difficult to use/easy to use

	Short		Long	
	Mean	SD	Mean	SD
HC	3.53	1.25	3.93	1.03
ES	4.07	0.96	4.07	0.88
PF	4.27	1.03	3.93	1.03

Question 1E, Overall – Rigid/flexible

	Short		Long	
	Mean	SD	Mean	SD
HC	2.60	0.51	2.07	1.03
ES	3.40	1.06	3.73	0.96
PF	3.93	1.22	3.73	1.03

Table 14 (continued)

Question 2D- Difficulty in searching for specific information

	Short		Long	
	Mean	SD	Mean	SD
HC	3.33	0.98	3.47	1.13
ES	3.67	0.98	4.27	0.88
PF	3.93	0.88	3.73	1.16

Question 2F - Frustration in searching for specific information

	Short		Long	
	Mean	SD	Mean	SD
HC	3.13	0.99	3.40	0.99
ES	3.47	0.99	3.87	0.74
PF	3.60	0.83	3.47	1.19

Question 6D- Difficulty in understand location within the mode

	Short		Long	
	Mean	SD	Mean	SD
HC	3.93	1.28	3.80	1.37
ES	3.80	1.32	3.87	1.13
PF	4.33	0.82	3.80	1.37

Question 6F - Frustration in understand location within the mode

	Short		Long	
	Mean	SD	Mean	SD
HC	4.00	1.36	4.07	1.03
ES	3.67	1.40	4.00	1.07
PF	4.13	0.92	3.73	1.22

Question 7D – Difficulty navigating within the mode

	Short		Long	
	Mean	SD	Mean	SD
HC	3.60	1.183	3.87	1.06
ES	4.07	1.033	4.40	0.63
PF	4.00	1.000	3.87	1.36

Table 14 (continued)

Question 7F - Frustration navigating within the mode

	Short		Long	
	Mean	SD	Mean	SD
HC	3.47	1.30	3.73	0.96
ES	3.93	1.03	4.20	0.86
PF	3.87	1.19	3.80	1.08

Question 8 – Value of the Product

	Short		Long	
	Mean	SD	Mean	SD
HC	3.47	0.92	3.80	1.01
ES	3.67	0.98	4.33	0.72
PF	3.73	0.96	3.87	1.19

Question 10D – Overall Difficulty in the search tasks

	Short		Long	
	Mean	SD	Mean	SD
HC	3.33	0.72	4.00	0.76
ES	3.53	0.99	3.80	0.94
PF	3.73	0.80	3.40	0.91

Question 10F – Overall Frustration in the search tasks

	Short		Long	
	Mean	SD	Mean	SD
HC	3.40	0.99	3.67	0.98
ES	3.47	0.99	3.60	0.63
PF	3.80	0.78	3.53	0.83

Table 15

Significance of presentation mode for SP and LP condtions for the End-of-Session Questionnaire

Questions	Short Pages	Long Pages
1A-Overall Reaction, Wonderful	0.084	0.192
1B-Overall Reaction, Satisfying	0.452	0.852
1C-Overall Reaction, Stimulating	0.004 ES=PF>HC	0.025 ES>HC, HC=PF, ES=PF
1D-Overall Reaction, Ease of Use	0.068	0.826
1E-Overall Reaction, Flexibilty	0.001 ES=PF>HC	0.0005 ES=PF>HC
2-Searching for Information, Difficulty	0.214	0.125
2-Searching for Information, Frustration	0.328	0.334
3-Searching Table of Contents, Helpful	0.962	0.946
3-Searching Table of Contents, Difficulty	0.746	0.625
3-Searching Table of Contents, Frustration	0.457	0.715
4-Searching Index, Helpful	0.124	0.188
4-Searching Index, Difficulty	0.011 PF>HC, ES=PF, HC=ES	0.114
4-Searching Index, Frustration	0.609	0.465
5-Browsing the text, Helpful	0.840	0.371
5-Browsing the text, Difficulty	0.694	0.332
5-Browsing the text, Frustration	0.649	0.393
6-Understanding location, Difficulty	0.331	0.987
6-Understanding location, Frustation	0.526	0.661
7-Navigating, Diffculty	0.182	0.343
7-Navigating, Frustration	0.381	0.422
8-Value of the mode	0.697	0.367
9-Difficulty of search tasks	0.488	0.085
9-Frustration of search tasks	0.299	0.911

Note : Number indicates obtained probability of results due to chance

length, better than the computer modes. Significant differences were obtained for the following questions.

Question 1: Overall Impression

The overall impression of three presentation modes differed significantly in only two of the five dimensions – dull/stimulating and rigid/flexible.

For both LP and SP, the HC mode was least stimulating. For SP, the ES and PF modes were significantly more stimulating than the HC mode. For LP, only the ES was more stimulating than the HC. The subjects found the ES and PF modes to be significantly more flexible than the HC mode for both SP and LP.

Question 4B: Ease of use of the Index

Searching using the index was significantly easier in the PF mode than the HC mode for SP but not LP. There were no significant differences between the electronic modes or between the HC and ES modes. The results parallel some of the performance results for individual tasks. For Tasks 2, 3, 4, subjects spent more time using the search tools with HC.

Analysis of Preference Questionnaire

At the conclusion of the last session, the subjects were administered a questionnaire in which they were asked to rank order the three modes along a number of dimensions. There were questions dealing with overall preference, ease of use and learning, ease of finding information, ease of reading the content of the mode, ease of understanding the location within the mode, ease of jumping from section to section, ease of using the different search tools, preference for regular use of the modes, and purchase preference. Appendix K contained a detailed description of the preference questions.

Table 16 shows the mean and standard deviation of the preference rankings for the presentation modes for each of the questions. In that table, higher numbers indicate greater preference.

The preference data, being inherently ordinal, was analyzed using the Friedman Analysis of Variance by ranks. Table 17 indicates significance was attained on 11 of 12 of the questions, with more questions being significant under LP than SP. More importantly for later reference, the ES mode was preferred over the HC mode overall and in terms of ease of use, ease of finding information, ease of jumping from one location to another, and ease of using the index. Significant differences were obtained for the following questions.

Question 1: Preference

For SP, the ES and PF modes were both significantly more preferred than the HC. However, for LP only ES was significantly preferred to HC.

Question 6: Ease of Use

For SP, the subjects ranked the ES and PF modes as significantly easier to use than the HC mode. However, for LP the ES mode was rated as easier to use than both the HC and PF modes. The PF results for Question 6 were consistent with the book metaphor hypothesis for LP in that the PF mode was rated as more difficult to use. However, inconsistent with the same hypothesis with SP, the PF mode was rated as easier to use than the HC mode.

Question 7: Ease of Learning

When subjects were asked about which media was easiest to learn, there were significant differences only with LP. With LP, the subjects ranked the HC mode as easier

Table 16

Means and Standard Deviations for the Preference Questionnaire [1]

Q1- Overall Preference

	Short		Long	
	Mean	SD	Mean	SD
HC	2.20	.41	1.67	.72
ES	2.40	.74	2.47	.64
PF	2.40	.63	1.80	.86

Q6-Ease of Use

	Short		Long	
	Mean	SD	Mean	SD
HC	2.60	.63	1.93	.70
ES	2.47	.74	2.67	.49
PF	2.60	.63	2.00	.93

Q7-Ease of Learning

	Short		Long	
	Mean	SD	Mean	SD
HC	2.67	.49	2.93	.26
ES	2.27	.80	2.20	.68
PF	2.00	.93	1.93	.70

Q8-Ease of Finding Information

	Short		Long	
	Mean	SD	Mean	SD
HC	1.73	.46	1.83	.74
ES	2.47	.83	2.80	.56
PF	2.73	.46	2.13	.74

Table 16 (continued)

Q9-Ease of Reading Content of Text

	Short		Long	
	Mean	SD	Mean	SD
HC	2.67	.62	2.80	.41
ES	2.33	.72	2.67	.49
PF	2.53	.64	2.33	.72

Q10-Ease of Location

	Short		Long	
	Mean	SD	Mean	SD
HC	2.73	.59	2.80	.56
ES	2.00	.85	1.93	.70
PF	2.33	.62	2.47	.64

Q11-Ease of Jumping from one location to another

	Short		Long	
	Mean	SD	Mean	SD
HC	1.60	.51	1.80	.68
ES	2.53	.74	2.73	.46
PF	2.67	.62	2.67	.49

Q12-Ease of the Table of Contents

	Short		Long	
	Mean	SD	Mean	SD
HC	2.38	.51	2.50	.53
ES	2.77	.44	2.30	.48
PF	2.77	.44	2.70	.48

Table 16 (continued)

Q13-Ease of Index

	Short		Long	
	Mean	SD	Mean	SD
HC	2.07	.26	2.00	.00
ES	2.93	.26	3.00	.00
PF	2.93	.26	3.00	.00

Q14-Ease of Browsing

	Short		Long	
	Mean	SD	Mean	SD
HC	1.67	1.00	2.07	.83
ES	2.33	.50	2.71	.61
PF	2.56	.53	1.86	.53

Q15-Regular Usage

	Short		Long	
	Mean	SD	Mean	SD
HC	1.87	.92	1.87	.74
ES	2.20	.77	2.73	.59
PF	2.27	.70	2.13	.83

Q16-Cost Analysis

	Short		Long	
	Mean	SD	Mean	SD
HC	1.80	.77	1.67	.72
ES	2.33	.72	2.60	.63
PF	2.53	.64	2.07	.80

Note. [1] Higher numbers indicate greater preference

Table 17

Significance of presentation mode for SP and LP condition for each question on Preference Questionnaire

Question	Short Pages	Long Pages
Q1-Overall	0.001 HC<ES=PF	0.038 HC<ES, HC=PF, ES>PF(.057)
Q6-Ease of Use	0.002 HC<ES=PF	0.042 HC=PF<ES
Q7-Ease of Learning	0.132	0.001 ES=PF<HC
Q8-Ease of finding Information	0.001 HC<ES=PF	0.008 HC=PF<ES
Q9-Ease of Reading Content of Text	0.347	0.043 PF<HC,HC=ES, ES=PF
Q10-Ease of Location	0.033 ES<HC, HC=PF, ES=PF	0.002 ES<HC=PF
Q11-Ease of Jumping from one Location to Another	0.000 HC<ES=PF	0.001 HC<ES=PF
Q12 - Ease of Table of Contents	0.103	0.607
Q13-Ease of Index	0.000 HC<ES=PF	0.000 HC<ES=PF
Q14-Ease of Browsing	0.177	0.023 PF<ES, HC=ES, HC=PF
Q15-Regular Usage	0.423	0.015 HC<ES, HC=PF, ES=PF
Q16-Cost Analysis	0.056 HC<PF,HC=ES, ES=PF	0.015 HC<ES, HC=PF, ES=PF

Note: Greater than (>)indicates a greater preference; Numbers indicate obtained probability of result due to chance.

to learn than the ES and PF modes. This is understandable because the subjects were familiar with use of HC prior to the start of the study. For SP, the subjects also ranked the HC as easier to learn than the computer modes but this difference was not significant.

Question 8: Ease of Finding the Information

The results of Question 8 parallel those for ease of use (Question 6). For SP, the ES and PF modes were rated as being easier to find the information than HC. However with LP, only the ES mode was rated higher than the HC mode. The HC and PF modes did not significantly differ from one another for LP.

Question 9: Ease of Reading Content of Text

Ease of finding the content of the text pertains to Stage 2. When the subjects were asked about the ease of reading the text, only the LP condition resulted in a significant presentation mode difference. For LP, the HC mode was significantly preferred over the PF mode. There was no statistically significant difference between the HC and ES modes. The HC-PF difference is what might be expected based upon the presumed difficulty in using the PF mode under the LP condition. For SP, the HC mode was preferred over both the ES and PF modes but this difference was not significant.

Question 10: Ease of Locating the Answer

For SP, the HC mode was deemed easier to find the location of the answer than the ES mode. There was no statistical difference between HC and PF modes and the ES and PF mode. However for LP, the HC and PF mode were both preferred over the ES mode in terms of ease of locating the answer.

Question 11: Ease of Jumping From One Location to the Next

For both LP and SP, the subjects found that that it was significantly easier to jump from one location to the next with the computer modes than the HC mode. There was no significant difference between the computer modes.

Question 13: Ease of Searching Through the Index

Ease of searching the index is one of the two questions which clearly assesses Stage 1. For both SP and LP, the ES and PF modes were deemed easier to search the index than HC. Once again, consistent with the task-by-task analysis and the end-of - session questionnaires, the subjects had difficulty with the HC index.

Question 14: Ease of browsing the text

Ease of browsing the text relates to Stage 2. Somewhat surprisingly, there was only one significant difference in the performance for the three modes in terms of ease of browsing. For LP, the PF mode was judged more difficult to browse than the ES mode.

Question 15: Preference for Regular Use

With LP, ES was preferred as the presentation mode that subjects would use regularly over the HC mode. There was no statistical difference between the HC and PF mode and ES and PF mode.

Question 16: Purchase Preference

For LP, assuming cost was not a factor, subjects preferred purchasing the ES mode over the HC mode. There was no statistical difference between the HC and PF modes and the ES and PF modes.

Manipulation Checks

The preference questionnaire also afforded the opportunity to check on the equivalence the hardcopy and computer modes in terms of the darkness and clarity of the font, the size of the characters, and the background contrast.

Question 2 : Darkness of the Type of font

When subjects were asked about a noticeable difference in the darkness of the type of font, the majority of subjects claimed they detected no difference. Only 6 (3 SP and 3 LP) of 30 subjects claimed they noticed a difference (See Table 16). Of those 6 subjects, half of them stated that the HC was darker than the computer and the other half claimed the opposite. However, all six of these subjects reported that this difference did not affect their performance. The six subjects who noticed a difference were asked to rate the degree to which they felt that this factor affected their performance on a 9 point scale (with 1 being not at all and 9 being greatly affected their performance). The means for both the hardcopy and computer modes were 1.67.

Table 18

Number of Subjects reporting differences in the Darkness of the font

	Page Length	
Presentation Mode	Short Pages	Long Pages
	Pages	
Hardcopy	2	1
Computer	1	2

Question 3 : Clarity of the Type of Font

When subjects were asked about a noticeable difference in the clarity of the type of font, 26 of 30 subjects claimed neither the computer nor the hardcopy differed in terms of the clarity (See Table 17) . Of the 4 subjects (3 LP, 1 SP) that reported a difference, three stated that the clarity of font was greater with the hardcopy mode than the computer modes. However, all four of these subjects reported that this difference did not affect their performance. Both the means for the hardcopy and computer modes were 2.00.

Table 19

Number of Subjects reporting differences in the Clarity of the type of font

	Page Length	
Presentation Mode	Short Pages	Long Pages
Hardcopy	1	2
Computer	0	1

Question 4: Size of the Characters

Twenty-five of thirty subjects did not detect any difference in the size of the characters between the hardcopy and computer modes. Of the five subjects (2 SP, 3 LP), four claimed that the size of the characters were greater with the hardcopy mode than the computer mode (See Table 18). The means for both the hardcopy and computer modes were 2.00.

Table 20

Number of Subjects reporting differences in the Size of the Characters

	Page Length	
Presentation Mode	Short Pages	Long Pages
Hardcopy	2	2
Computer	0	1

Question 5: Background contrast

When subjects were asked about any noticeable difference in the background contrast, 24 of 30 subjects detected no difference between the two modes. Of the 6 (4 SP, 2 LP) subjects reporting differences, four stated that the computer contained the greater background contrast (See Table 19). The mean for the hardcopy mode was slightly higher at 3.5 than the mean for the hardcopy mode which was 1.5.

Table 21

Number of Subjects reporting differences in the Background Contrast

	Page Length	
Presentation Mode	Short Pages	Long Pages
Hardcopy	1	1
Computer	3	1

CHAPTER 4

DISCUSSION

Three major results emerged from the present study. First, there appeared to be no differences in the time it took the subjects to formulate the problem (FTime), time to search the index (S1Time), time to browse the document (S2Time), and total task time(TskTime) as function of presentation mode for either SP and LP. Second, the only significant performance difference was in the percent correct measure where the PF mode resulted in the poorest performance. Clearly these results do not support either the amount of text displayed hypothesis or the book metaphor hypothesis. Third, despite the failure to find the differences with the performance measures, significant differences were obtained with the subjective data showing a preference for the ES mode over the HC mode.

The book metaphor hypothesis predicted that there should be a significant mode effect in S2Time for both SP and LP. Given that significance in S2Time was not obtained for either SP or LP, the overall results are more consistent with amount of text displayed hypothesis than the book metaphor hypothesis. As predicted by amount of text displayed hypothesis, there were no significant differences in presentation mode for SP. However, contrary to predictions derived from amount of text displayed hypothesis, the expected mode differences did not emerge for LP.

While specific predictions were not upheld for the amount of text displayed hypothesis, perhaps it was supported in a general sense. The HC text was made to resemble an electronic version to control for some of the differences in the amount of text displayed. Unlike a real book, the pages in the HC mode were not back-to-back. Rather than 92 lines (2 pages back-to-back) of text being in view of the subject at any one point in time, there were only 46 lines in the HC-LP document compared to 23 lines in the electronic version. Thus, a 4 to 1 difference (92 lines versus 23 lines) was reduced to a 2 to 1 difference (46 lines versus 23 lines) and this may not have been sufficient to result in time differences between the HC and the electronic modes. Secondly unlike a true HC index, which is a full page, back-to-back (92 lines), and sometimes in two columns per page, the HC index in the present study was made identical to what was viewed on the CRT in the electronic versions (23 lines). Thus, the difference in the amount of text displayed for the HC and the electronic versions was greatly reduced when compared to the difference between a real book and the electronic versions. Given this reduction in the amount of text displayed, it may have been unreasonable to expect a difference between the HC and electronic versions.

On the one hand, these results add to the growing body of literature that shows no differences between the HC and computer presentation of the same text material when one controls relevant variables. It is important to remember that Gould et al. (1987b) found that the HC-CRT differences disappeared when the physical aspects of the HC-CRT presentation were controlled. Even though the dependent measure in the present experiment was search time and not reading speed as in Gould et al. (1987b) studies, HC-CRT performance differences were not present when the HC presentation was made

similar to the information displayed on the CRT. The results are also consistent with Richardson et al. (1988) who found that the number of lines of text on a computer screen did not affect differences. Taken together, these results (Gould et al. 1987b and Richardson et al., 1988) and those of the present study suggest that when these factors are controlled, there may be no inherent difference in the presentation of text in HC versus the CRT. On the other hand, there are other explanations for the present results (See next section).

Counteracting Effects of User Experience and HC Difficulty

The primary explanation for the failure to obtain significant time differences among the three modes is rooted in a counteracting of two factors – the experience level of the subjects and the difficulty of utilizing the HC mode. There were a number of differences between the Biers et al. (1995) usability study and the present study that provide insight into the lack of significant differences with regard to the performance data.

First, the level of experience differed between the Biers et al. (1995) usability study and the present study. Sixty-seven percent of the subjects in the Biers et al. (1995) study were beginners and 33% of the subjects were intermediates. There were no subjects classified as advanced computer users. Closer inspection of the Biers et al. (1995) data revealed that only the beginner computer subjects found the information significantly faster with the HC mode than the computer modes. However, in the present study, 17% were beginners, 40% were intermediate, and 43% were advanced users. Since the present study had subjects that were more proficient and experienced with computers,

the differences in the use of HC and the computer modes could be expected to be reduced.

Second, in the present study, the level of experience significantly differed for those subjects in the SP and LP condition. For SP, the subjects were equally split among the three experience levels whereas 67% of the subjects in the LP condition were classified as advanced computer users. Since differences were expected under both hypotheses for LP, the experience factor may have worked against finding significance under LP. The greater the experience level of the subjects, the better the proficiency in using the computer tools. Thus, the utilization of more experienced users would probably result in reduced time differences between the HC and computer modes.

Support for the experience explanation comes from several post hoc analyses. First, subject experience was correlated with the four time-based measures. Experience was negatively correlated with time for F1Time (ρ=-0.337, p=0.001), S1Time (ρ=-0.259, p=0.014), and TskTime (ρ=-0.207, p=0.050). That is, the more experienced users took less time to formulate, to search the index, and to find the correct answer. It is clear from this analysis that experience level is related to performance on FTime, S1Tme, TskTime, but not in S2Time as was expected.

To see if the experience factor could have mitigated the expected outcome, a 2 (page length) X 3(presentation mode) X 3(experience level) mixed ANOVA was performed on the data. For two of the measures, S1Time (F=2.35, p=0.062) and TskTime (F=2.16, p=0.082), the interaction of presentation mode and experienced approached significance. Figure 10 presents the mode by experience interaction for overall task time. Inspection of this figure shows that with beginner computer users, the

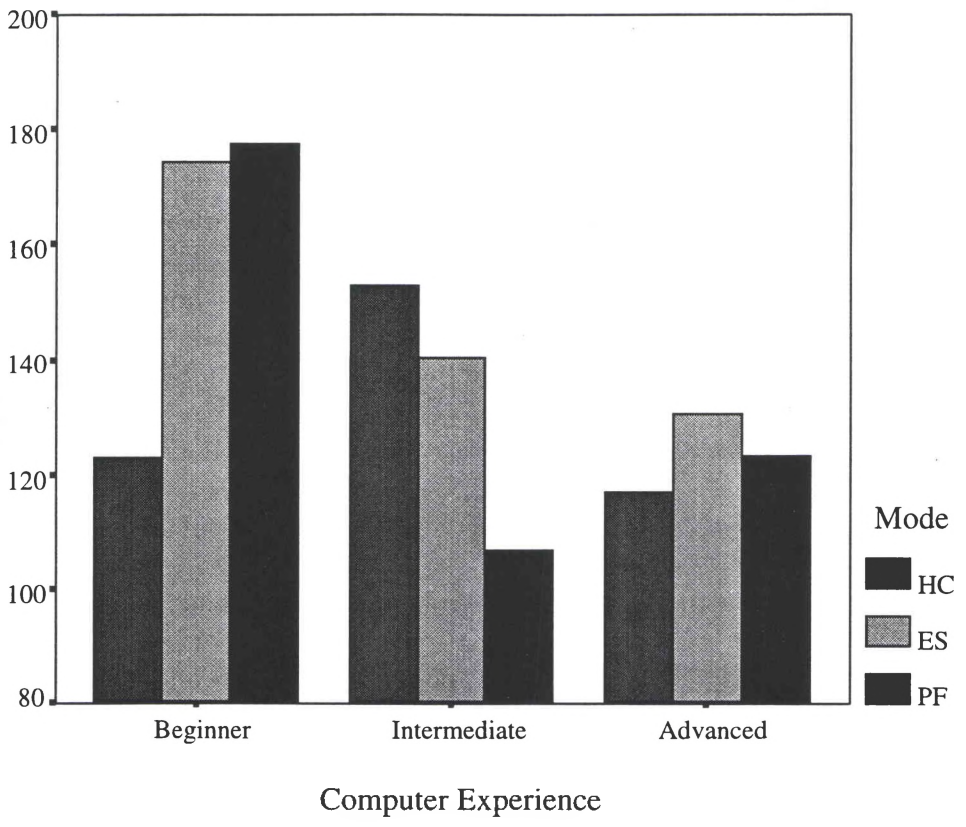


Figure 10. Mode by Experience interaction for overall task time

HC took less total time than either the ES or PF modes whereas with advanced computer users, there was little or no difference between the three modes with overall task time. Therefore, if one only looks at the beginner subjects, the results are consistent with the Biers et al. (1995) usability study. Given that most of the subjects in the present study were intermediate and advanced computer users, the expected difference may have been obscured, particularly for LP where most of the users were advanced. This evidence suggests that if the subject pool was limited to beginners, the predicted significant effects might have been attained.

Figure 11 presents the mode by experience interaction for S1Time (F=2.35, p=0.062). Like TskTime, beginner subjects took less time with the HC search tools than with the ES and PF modes. With advanced subjects, the differences among the three modes were less than with beginner subjects, which parallels the results of overall task time. Instead of finding significant differences in S2Time as predicted, the results approached significance for S1Time.

Taken together, these results have two important implications. First, these results imply that the amount of text displayed hypothesis might have been upheld had less experienced subjects been employed. Second, these results suggest that the locus of the effect was not in S2Time, but may have been in S1Time. This is supported by some of the task-by-task performance analyses which revealed significant differences in S1Time. The hypotheses in this study were based upon the results of the Biers et al. (1995) study which only measured overall task time. Based upon direct behavioral observation, Biers et al. (1995) concluded that a major factor responsible for slower search times using the electronics text version was the difficulty in locating information in the document

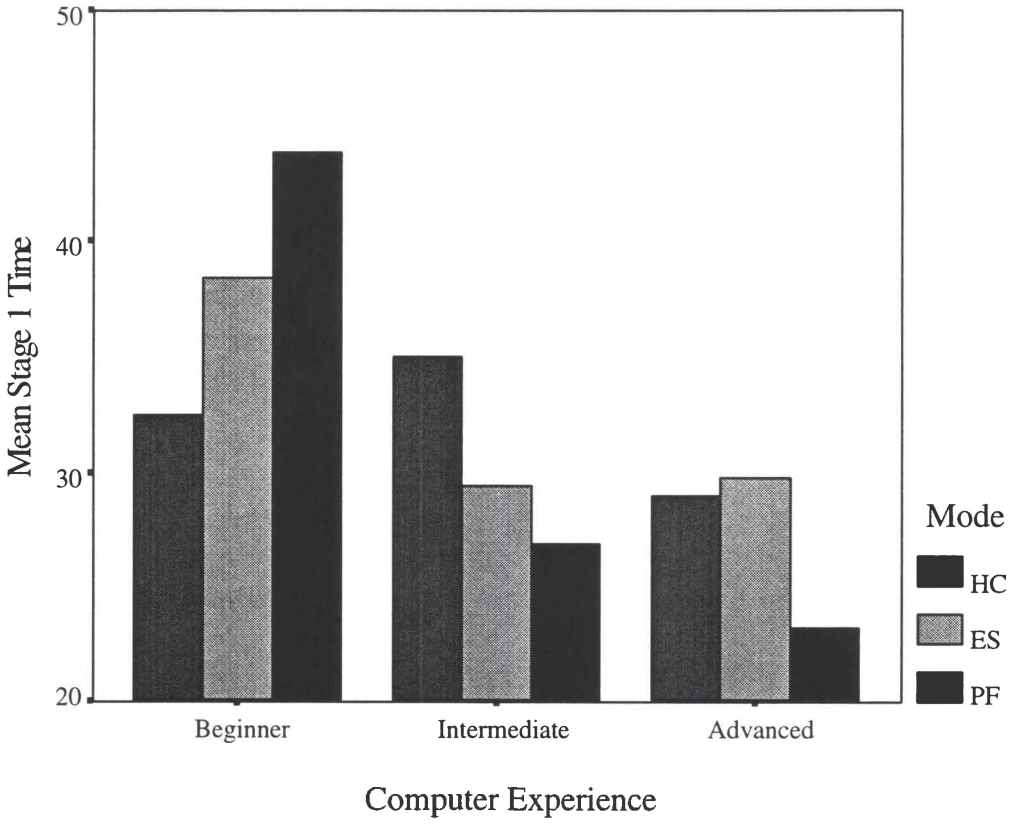


Figure 11. Mode by Experience interaction for S1Time

itself (Stage 2). However, users in the Biers et al. (1995) study also had difficulty using the computer search tools (Stage 1). Perhaps Stage 1 was the source of the overall task time differences in the Biers et al. (1995) study and not in Stage 2 as was predicted. Given that Biers et al. (1995) did not measure the time spent in the different stages, it was impossible to isolate the locus of the effect. One of the methodological improvements in the present study was the separation of total task time into three stages. The present results indicated little evidence in showing mode differences in S2Time. The aforementioned post hoc analyses with experience suggest there may have been a mode effect in Stage 1, not in Stage 2.

Experience alone is probably not sufficient to explain the lack of significant differences. Another factor that probably contributed to the failure to obtain significant time differences was the difficulty in using the HC mode itself. For the purposes of experimental control, the HC mode was made similar to the computer modes in terms of the amount of text displayed to the subject. In the HC versions, the text was not printed on back-to-back pages and in the SP condition, the text was limited to 23 lines rather than 46 lines. In addition, the index of the HC presentation mode was also limited to 23 lines per page. This greatly increased the thickness of the book resulting in a HC-SP document and its index containing over 380 pages. For LP, the HC presentation mode contained over 220 pages, and this is much greater than would have been the case with a real book. This made these HC documents awkward, difficult, and cumbersome for the subject to utilize. Thus, the subjects were forced to leaf and search through a large stack of single-sided paper which may have resulted in longer task times than would have been the case

for a true book which would have contained approximately 100 pages(46 lines per page, back-to-back pages with a full page table of contents and index).

Thus, it may not have been solely the ease of using the computer for advanced subjects which resulted in the non-significant presentation mode effect, but the difficulty in using the HC version in conjunction with more proficient computer users. To verify that the HC book was indeed more difficult to search than a regular book, an additional control condition would have to be run in which the HC was printed back-to-back with both the document and index being 46 lines per page. If differences were found with the current HC document and a HC document designed to resemble a real book, this would support the amount of text displayed hypothesis.

Methodological Explanations

There are two additional reasons for the failure to find differences in the time-based measures. The first explanation involves the variability inherent in the time-based measures themselves. Time-based measures, particularly completion time, notoriously result in large within-cell variability and thus contribute to error variability. With large error variability, it is much more difficult to detect differences that might have been present.

The problem of the variability of time-based measures was exacerbated by having an extremely liberal time limit. A ten-minute time limit was allotted per question which was much more time than most subjects needed. The time to locate the correct answer varied from 13 to 589 seconds with the median time being 99 seconds. Nine-two percent of the subjects found the answers in less than five minutes, which indicates that the 10 minute time limit was more than ample. With no time pressure, some subjects could take

as much time as necessary. If the subjects did not find the answer immediately, they had ample opportunity to find the correct answer within the allotted time. In the real world, most people would probably give up after trying one or two times to find the answer and probably never would have spent six minutes looking for the answer, let alone 10 minutes. This liberal time element could have contributed to the large within-cell variability in the time measures. For a future study, perhaps subjects should be put under a time pressure so the sensitivity to detect differences would lie in the correctness measure.

Another factor that may have contributed to the variability in the time-based measures is the measurement error associated with the experimenter having to control the timing. Rather than automatically controlling the timing, the experimenter had to make judgments in real time as to which key to depress (i.e. the stage) and had to respond quickly to any stage shifts. Any lag in the experimenter responding would contribute to measurement error. Therefore, since the time-based stage measures were not independent, any inaccuracies in one stage would automatically lead to inaccuracies in measurement of another stage. A methodological improvement would be to capture the times directly off the subject's computer based upon the length of time various list boxes were open and to videotape the subjects during the hardcopy presentation mode.

The second factor which may have contributed to the failure to find performance differences is the ease of the search queries in conjunction with the nature of the sourcebook itself . The subjects answered the queries with an average of 95% correct and perceived the difficulty of the tasks to be fairly easy (3.75 on a five point scale with higher numbers indicating less difficulty). This suggests that the tasks may have been too

easy or that the sourcebook was organized in such a fashion that the information was easy to find. In comparison to the sourcebook used in the Biers et al. usability study (1995), the answers to the questions in the 1996 Your Federal Income Tax for Individuals-Publication #17 were directly found in a few contiguous lines. Once located, the answer was obvious. In the Biers et. al usability (1995) study, the answer to some of the questions involved integrating the material over several pages and the answer was not as obvious to the subject.

Page Flipping Mode and Design Implications

The only significant result found in the performance data was that the PF mode for LP resulted in a lower percentage correct than the other two presentation modes. This suggests that this particular mode, in accordance with the book metaphor hypothesis, was awkward and difficult to use. With the PF mode for LP, the subjects had to engage in two physically different actions to view the text –the use of the vertical scrollbar to move within a page and the use of the backward/forward buttons to retrieve another page. Not being able to view the entire page at one time is not faithful to book metaphor.

The difficulty of using the PF mode for LP is supported by the preference data. The subjects rated the PF mode for LP lower than the other modes in terms of ease of learning, ease of reading the content or text, ease of browsing, and ease in use of finding the information. These results indicate that the subjects had trouble using and finding information using the LP-PF mode. From a design standpoint, this indicates that the page length should be limited to the number of lines that can be displayed on the computer screen. The fact that the PF mode was rated higher in terms of finding information with SP supports this contention.

Subjective Data

Despite the lack of performance differences, there were significant differences in preference for the three modes. This is consistent with previous research (e.g. Schwarz et al. (1983), Gould et al. (1987b), and Dillon et al. (1989, 1990)) which has shown that despite the lack of performance differences, there were significant differences in preference.

Overall, for both SP and LP in the present study, the subjects preferred the ES mode to the HC mode. In addition, the ES mode was preferred over the HC mode in terms of the ease of use, ease of finding the information, ease of jumping from one location to another, and ease of using the index. These results are consistent with the explanation based upon the interaction of the subject's computer experience and the difficulty of using the HC condition. However, rather than counteracting one another as with performance, these factors acted in an additive fashion to make the differences more apparent with the preference data. In terms of preference, the positive effect of the ease of using the computer modes (i.e. greater computer experience) combined with the negative effect of the difficulty of using the HC mode to result in a clear preference of ES over HC.

For the remainder of the preference questions (ease of learning, ease of reading the content of the text, and ease of location), significant differences between the modes were only obtained under the LP condition. Given that the subjects were more experienced under LP than SP, the ES-HC differences would be expected to be larger under LP.

There are some interesting contrasts between the preference data in the Biers et al. usability study (1995) and those of the present study which support an experience interpretation. Consistent with the present study, there was an overall preference for the electronic version (i.e. ES) over HC which was manifested for both beginner and intermediate users. However, there were more Electronic/HC book differences which were significant for the intermediate users than for beginners. For instance, the electronic book was more preferred in terms of the ease of use and ease of finding information for intermediate users, but not for beginners. In the current study, consistent with expectations based upon more experienced users, these same questions resulted in a preference for the ES mode over the HC mode.

There were some questions that showed a preference for the HC mode compared to the computer modes. For ease of learning and understanding their location within the book, subjects preferred the HC to the computer modes. It was expected that subjects would find it easier to learn and understand their location within the book under the HC mode because of their long history of using books.

Future Research

Two additional studies need to be conducted to shed light on the interpretation of this present study. First as previously mentioned, a traditional HC book condition needs to be run to provide a baseline for assessing the performance of all modes. This condition should resemble a typical book— 46 lines per page, back-to-back pages, with a full-page index and table of contents. The performance of users utilizing the traditional book would shed light on the difficulty of using the HC version in the present study and possibly provide support for the viability of the amount of text displayed hypothesis.

Secondly, the experience explanation should be tested systematically. A study should be conducted in which user experience is factorially combined with the variables in the present study.

APPENDIX A

Search Queries

1. You have bought real estate property during the year (1996). When does the seller of the real estate quit paying taxes on that property?
2. Does the Veterans of Foreign Wars (VFW) qualify as a charitable organization?
3. Is there any limitation on the amount of itemized deductions you want to deduct (a dollar amount)?
4. Your husband is in a nursing home for the purposes of medical care and you pay for the entire cost. Can you deduct this on your tax return?
5. I refinanced my home in 1996 and paid "points". Can I deduct the entire amount as interest on my 1996 return?
6. What are the requirements that must be met in order to deduct educational expenses?
7. You bought a sofa 3 years ago (1993) for $1000. In April of 1996, the sofa was damaged in a fire. You estimate that it would cost $1300, to replace it. If you had sold the sofa before the fire you could have gotten $600. What is the amount of the loss, $1300 or $600 and why?
8. Name 3 criteria that must be met in order to claim a dependent for the purposes of a medical deduction?

9. How long does a business trip need to last before an employee can deduct the cost of meals?
10. What 3 things must you be able to show the IRS to take a deduction for a casualty loss?
11. You sold your home on August 1, 1996. Can you deduct the allowable home mortgage interest for the month of August?
12. You are an auto mechanic. With everything becoming computerized in the car, you take a special course to gain expertise in this area. Can you deduct this course?
13. Name 3 taxes that are not deductible as real estate taxes?
14. What are the standard mileage rates for 1996?
15. You have an eye condition that will not improve. You wish to take a higher standard deduction for partially losing your sight. What can you do to bypass attaching a certified statement to your tax form?

APPENDIX B

Hardcopy Questionnaire

Name of Participant: Date:

Interviewer: Experience Level:

Please evaluate the Hardcopy (Paper) format without considering the content quality of "1996 Federal Income Tax for Individuals-Publication 17", unless a specific question regarding the content is asked.

1. Overall reactions to the product:

terrible				wonderful
1	2	3	4	5
frustrating				satisfying
1	2	3	4	5
dull				stimulating
1	2	3	4	5
difficult				easy
1	2	3	4	5
rigid				flexible
1	2	3	4	5

2. Searching for specific information is:

difficult				easy
1	2	3	4	5
frustrating				not at all frustrating
1	2	3	4	5

3. Searching using the table of contents is:

not at all helpful very helpful
1 2 3 4 5

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

4. Searching using the index is:

not at all helpful very helpful
1 2 3 4 5

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

5. Searching browsing the text is:

not at all helpful very helpful
1 2 3 4 5

difficult easy
1 2 3 4 5

6. Understanding your location within the product is:

difficult easy
1 2 3 4 5

7. Navigating within the product is:

difficult easy
1 2 3 4 5

8. To me, the product is:

not at all valuable very valuable
1 2 3 4 5

9. Specifically, finding the answer to the questions listed below was:

Question #1

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #2

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #3

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #4

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #5

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

10. If you could add or change any features to this product, what would they be?

APPENDIX C

Electronic Scrolling Questionnaire

Name of Participant: Date:

Interviewer: Experience Level:

Please evaluate the Electronic Scrolling format without considering the content quality of "1996 Federal Income Tax for Individuals-Publication 17", unless a specific question regarding the content is asked.

1. Overall reactions to the product:

terrible					wonderful
1	2	3	4	5	
frustrating					satisfying
1	2	3	4	5	
dull					stimulating
1	2	3	4	5	
difficult					easy
1	2	3	4	5	
rigid					flexible
1	2	3	4	5	

2. Searching for specific information is:

difficult					easy
1	2	3	4	5	
frustrating					not at all frustrating
1	2	3	4	5	

3. Searching using the table of contents is:

not at all helpful very helpful
1 2 3 4 5

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

4. Searching using the index is:

not at all helpful very helpful
1 2 3 4 5

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

5. Searching browsing the text is:

not at all helpful very helpful
1 2 3 4 5

difficult easy
1 2 3 4 5

6. Understanding your location within the product is:

difficult easy
1 2 3 4 5

7. Navigating within the product is:

difficult easy
1 2 3 4 5

8. To me, the product is:

not at all valuable very valuable
1 2 3 4 5

9. Specifically, finding the answer to the questions listed below was:

Question #1

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #2

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #3

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #4

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #5

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

10. If you could add or change any features to this product, what would they be?

APPENDIX D

Electronic Page Flipping Questionnaire

Name of Participant: Date:

Interviewer: Experience Level:

Please evaluate the Electronic Page Flipping format without considering the content quality of "1996 Federal Income Tax for Individuals-Publication 17", unless a specific question regarding the content is asked.

1. Overall reactions to the product:

terrible				wonderful
1	2	3	4	5
frustrating				satisfying
1	2	3	4	5
dull				stimulating
1	2	3	4	5
difficult				easy
1	2	3	4	5
rigid				flexible
1	2	3	4	5

2. Searching for specific information is:

difficult				easy
1	2	3	4	5
frustrating			not at all	frustrating
1	2	3	4	5

3. Searching using the table of contents is:

not at all helpful very helpful
1 2 3 4 5

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

4. Searching using the index is:

not at all helpful very helpful
1 2 3 4 5

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

5. Searching browsing the text is:

not at all helpful very helpful
1 2 3 4 5

difficult easy
1 2 3 4 5

6. Understanding your location within the product is:

difficult easy
1 2 3 4 5

7. Navigating within the product is:

difficult easy
1 2 3 4 5

8. To me, the product is:

not at all valuable very valuable
1 2 3 4 5

9. Specifically, finding the answer to the questions listed below was:

Question #1

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #2

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #3

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #4

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

Question #5

difficult easy
1 2 3 4 5

frustrating not at all frustrating
1 2 3 4 5

10. If you could add or change any features to this product, what would they be?

APPENDIX E

Last Session: Preference Questionnaire

Interview Format

Name of Participant: Date:

Rank Order the 3 Media.

(Note to evaluator: Put "1" for most preferred; "3" for least preferred. No need to ask why but if they offer, let them go)

1. Overall, which media did you most prefer? Least prefer? WHY?

_____Hard Copy Documentation (1996 Federal Income Tax Guide)
_____Electronic Scrolling
_____Electronic Page Flipping

The Following Questions deal with the Hardcopy versus Electronic version.

Did you notice any difference between the electronic versions and the hardcopy in terms of:

2. Darkness of the type of font

___Yes : Which was darker?

___Computer

___Hardcopy

How did this difference affected performance?

Not at all 1 2 3 4 5 6 7 8 9 **greatly**

___No

3) Clarity of the type of font

___Yes : Which had more clarity?

___Computer

___Hardcopy

How did this difference affected performance?

Not at all 1 2 3 4 5 6 7 8 9 **greatly**

___No

4) Size of the characters

___Yes : The size of the characters were greater in which media?

___Computer

___Hardcopy

How did this difference affected performance?

Not at all 1 2 3 4 5 6 7 8 9 **greatly**

___No

5) The background contrast

___Yes : The background contrast was greater in which media?

___Computer

___Hardcopy

How did this difference affected performance?

Not at all 1 2 3 4 5 6 7 8 9 **greatly**

___No

I would now like you to further evaluate the three tools **for finding information** which you utilized over the last three sessions--Hard Copy documentation, Electronic Scroll and the Electronic Page Flipping interface). In the past, when you purchased books at the library or bookstore, the information was in its hard copy format. The question is: how did each of the media affect your ability to locate information? I would now like to walk you through the following questionnaire, comparing the Electronic Scrolling and Page Flipping Interface to hardcopy.

Do you have any questions?

Please rank the three interfaces along the following criteria.

(For the evaluator only!!)*"1" means that you considered this interface the easiest to learn. "2" means that you considered the interface neither easy nor difficult to learn. And "3" means you considered this interface the most difficult to learn. If you liked any two interfaces equally, mark both of the interfaces with "1"s and the one you liked the least with a "2".*

6. Which media is easier to use? WHY?

_____ Hard-copy book
_____ Electronic Scrolling
_____ Electronic Page Flipping

7. Which media is easiest to learn? WHY?

_____ Hard-copy book
_____ Electronic Scrolling
_____ Electronic Page Flipping

8. In which media is it easiest to find information? WHY?

_____ Hard-copy book
_____ Electronic Scrolling
_____ Electronic Page Flipping

9. In which media is it easiest to read the content/text of the book? WHY?

_____ Hard-copy book
_____ Electronic Scrolling
_____ Electronic Page Flipping

10. In which media is it easiest to understand your location within the book? WHY?

_____ Hard-copy book
_____ Electronic Scrolling
_____ Electronic Page Flipping

11. In which media is it easiest to jump from one location to the next? WHY?

_____ Hard-copy book
_____ Electronic Scrolling
_____ Electronic Page Flipping

Next, you will be asked to compare the electronic and Hardcopy features.

Please use the same scale (1-3) which you used above.

12. With which media is it easier to search through the table of contents? WHY?

_____Hardcopy Documentation
_____Electronic Scrolling
_____Electronic Page Flipping

13. With which media is it easier to search through the index? WHY?

_____Hardcopy Documentation
_____Electronic Scrolling
_____Electronic Page Flipping

WHY?

14. With which media is it easier to browse the text to search for answers? WHY?

_____Hardcopy Documentation
_____Electronic Scrolling
_____Electronic Page Flipping

15. Which media would you be most likely to use regularly? WHY?

_____ Hard-copy book
_____ Electronic Scrolling
_____ Electronic Page Flipping

16. Assuming cost is not a factor, which media would you be most likely to buy? WHY?

_____ Hard-copy book
_____ Electronic Scrolling
_____ Electronic Page Flipping

APPENDIX F

Participant Consent Form

Description and Duration of Experiment

I am here to ask for your participation in a research study that I am conducting with Dr. David Biers in the Department of Psychology at the University of Dayton.

This experiment involves searching for information using a tax reference guide. This experiment is composed of three sessions, each of which lasts approximately 1 hour and 15 minutes. You are free to choose not to participate or stop participating in this experiment at any time if you should desire.

Confidentiality of Data:

The information that you provide will remain completely confidential, will be securely stored, and will only be viewed by the researchers involved in this study. After completing this form the experimenter will immediately place it in a folder in order to ensure the confidentiality of your name. Thus, your name will not be associated with any of your responses. Your data will be used collectively with all of the other participants' data.

Your participation or nonparticipation will have no effect upon your grades, or your relationship with your instructor, your educational institution, or the University of Dayton. In addition, there are no risks associated with your participation in this study. If

3) your wish to receive a summary of the results of this study, please print your full name and address on this consent form. If you have any questions, please do not hesitate to ask. Thank you very much for your cooperation and participation.

Consent to Participate

I have voluntarily decided to participate in this experiment. The experimenter has adequately answered all of my questions. I understand that I can voluntarily terminate my participation in this experiment at any time.

Signature_____ Date _____

APPENDIX G

User Profile and Experience Questionnaire

Name of Participant: Date:

If you do not understand any part of the questionnaire, please ask the experimenter for clarification.

1 Which category best describes your education level?

_____High School Graduate
_____Technical Degree
_____College Graduate
_____Post-College Graduate

2. How long have you been using a personal computer?

_____Never used a computer (If you checked this, see the experimenter)
_____Less than one year
_____1 - 2 years
_____3 - 5 years
_____6 - 7 years
_____8 - 10 years
_____Over 10 years
_____Don't know

3. How frequently do you use a computer at home?

_____Don't have a computer at home
_____Less than once a month
_____Once a month
_____Once a week
_____A few times a week
_____Once a day
_____More than once a day

4. How frequently do you use a computer at work?

_____Don't have a computer at work
_____Less than once a month
_____Once a month
_____Once a week
_____A few times a week
_____Once a day
_____More than once a day

5. Have you used Microsoft Windows (MS-Windows)?

_____Yes
_____No (See Experimenter)

6. How long have you been using Microsoft Windows?

_____Less than 6 months
_____6 months to 1 year
_____1 -2 years
_____3 - 4 years
_____Over 5 years

7. How comfortable are you with using Microsoft Windows?

_____Very comfortable
_____Somewhat comfortable
_____Not very comfortable
_____Not at all comfortable

8. I feel comfortable when I use any of the following: the File Manager, Control Panel for Windows 3.x, **OR** Windows Explorer and My Computer for Windows 95:

Yes No

9. I have run multiple Windows applications at the same time switching back and forth between applications when necessary:

Yes No

10. When I am using multiple Windows applications, I frequently move or re-size the windows

Yes No

11. I can copy and paste from one application to another

Yes No

12. I have used the right click option on the mouse to help me

Yes No

13. Imagine that you have purchased a new personal computer. Rank order the following in terms of your preference for **learning about your new computer** (1= most preferred and 5 = least preferred). (Note: Only use each number ONCE!)

_____Referring to hard-copy manuals as needed
_____Read hard-copy manuals (i.e. Getting Started or Introduction) from cover-to-cover
_____Trial and error
_____On-line help when needed
_____Interactive tutorials (a 4 minute multimedia explanation of how to use the system)

14. Imagine now that you have been using your computer for a couple of months. Rank order the following in terms of your preference for finding information when you have a computer problem (1 = most preferred and 6 = least preferred). (Note: Only use each number ONCE!)

_____Hard-copy documentation
_____On-line help
_____Trial and error
_____Consulting a knowledgeable friend or co-worker
_____Customer service/technical support
_____Interactive tutorials (a 4 minute multimedia explanation of how to use the system)

15. Reading textual information on screen (full text material that is on-screen) is something I:

dislike 1 2 3 4 5 like

16. Reading help feature is something I:

dislike 1 2 3 4 5 like

17. How often do you use a book (paper) to find information about your federal income tax form as opposed to searching for the information online?

_____Not applicable because I never prepared a tax return
_____Never
_____Less than once a month
_____Once a month
_____Once a week
_____Daily

APPENDIX H

General Instructions (First Session ONLY)

Thank you for agreeing to participate in the study. This project involves evaluating three different formats for finding information about jobs..

If at any point you don't under-stand the instructions or have questions, do not hesitate to stop me and ask.

Let me give you an overview of what I am going to ask you to do. Imagine that you are going to graduate or have graduated from college. You have decided that you want to find a job. Since you have been pre-occupied with finishing school, you have put little time into your job search. Where are you going to go to find information to answer your questions?

There are two basic approaches to finding information about the strategies for finding a job. The first is to use hard-copy documentation; i.e., a reference manual, getting-started documentation, or a book that you purchased at a bookstore. To find information, you browse or leaf through the book, use the table of contents, or go to the index. The second approach is to find this information in electronic format. Instead of buying or checking out a hardcopy book from the bookstore, you might get an electronic version of the book, like a Compact Disc. This approach is to use an electronic vertical

scrollbar to browse through a book to find information. The third approach-an electronic page flipping interface-- is relatively new. With the electronic page flipping interface, you can use a backwards and forwards buttons, like Netscape, the turn the pages.

Over the three sessions in which you have agreed to participate, we will be evaluating and comparing each of these three tools-- hard-copy books, electronic scrolling , and electronic page flipping. We will be using these three tools to find out information about jobs. Each individual sessions should last around an hour and a half.

On each day you will be doing several things. First, I will give you a chance to explore each of the interfaces--either the electronic scroll, the hard-copy documentation, or electronic page flipping. Then, I will ask you to use the tool to accomplish a series of specific tasks involving locating information. Then I will ask you to complete a questionnaire involving your subjective impression of the tool. Finally I will ask your some open-ended questions involving your general impression of the tool. This procedure will be repeated each session.

Do you have any questions?

Please read and sign the following Release Form.

USER COMPLETES RELEASE FORM

Before we start the actual study, I would like you to complete the following background questionnaire. This questionnaire requests additional information about your use of the computer and about your preference for the tools for finding information about jobs.

USER COMPLETES BACKGROUND QUESTIONNAIRE

APPENDIX I

Hardcopy Instructions

1 "Introduction to Hard Copy Instructions"

Today you are going to be using hard copy documentation (paper) to search for information. You are certainly familiar with how to browse and find information in a book.

There are three different ways to search for information using the book. First, you can go directly to the Table of Contents to find the appropriate general subsection which may contain the information. Alternatively, you can leaf through the pages to locate the information. Finally, you can utilize the alphabetical index of terms found in the back of the book.

2. User Explores

That's sufficient information to get you started. Now I will give you approximately 10 minutes to explore the paper version of the book on your own. Remember that your task is to use the table of contents, the index, or any other means to locate the answers to the search questions. You may explore any chapter you wish.

(USER EXPLORES FOR APPROXIMATELY 10 MIN)

(FIRST SESSION ONLY)

3. User Accomplishes Basic Tasks using Hard Copy Documentation

Now that you have had a chance to explore the "1996 Federal Income Tax for Individuals-Publication 17", I would like you to accomplish a series of basic tasks to find information about your federal income tax form.

Before you start the search tasks, I need to mention two additional points. First, before you begin the tasks of searching for the answers to the search questions, I would like you announce when you are ready to begin the search task. You may announce "Ready to Begin". At this time, I will initiate the timer and you will have ten minutes to search for the answers.

Secondly, when you think you have found the answer, I would like to you to announce aloud "I have found the answer" or words to that effect, and point to the answer on the paper.

If the answer falls over multiple pages, you will need to point to the entire answer by flipping through the pages. The experimenter will be seated the right and slightly behind you.

When you think you have found the answer, you will announce "I have found the answer" or words to that effect, and then physically point to the beginning of the passage that that contains the first sentence of the correct answer. If the experimenter has determined that the answer is correct, then he will stop the timer and ask you to read the next search question. If the experimenter determined that answer is incorrect then he will tell you that the answer is incorrect and to continue looking until the experimenter tells you to stop.

After the ten minutes are up, the experimenter will tell you to stop and ask when you are ready to begin the next task. After reading the next search question, the experimenter will ask if you are ready to begin the next task. When you are ready you will announce "Begin" and you will have another ten minutes to locate the correct answer.

By the two of us working together I can better understand how individuals utilize this tool.

(Second and Third Sessions)

Remember, as in the previous session(s) I would like you to announce when you have arrived at an answer and to point to it. If the answer falls along several pages, the subject must show the experimenter every pages in which the answer is contained.

(all Sessions)

Do you have any questions?

(USER PERFORMS TASKS)

4. User Completes Hard Copy General Questionnaire

Now I would like you to complete this general questionnaire regarding your subjective

impression of using hard copy version of the "1996 Federal Income Tax for Individuals-Publication 17" to find out information about the strategies for finding a information about how to complete your federal income tax form. In making your evaluation I would like you to focus on the hard copy documentation concept (paper), rather than to the specific contents of the "1996 Federal Income Tax for Individuals-Publication 17".

5. Session Debrief

Now I would like you to answer some open ended questions regarding use of hard copy documentation such as the "1996 Federal Income Tax for Individuals-Publication 17," to find out information about how to complete your federal income tax form.

APPENDIX J

Electronic Scrolling Instructions

1 **"Introduction to Electronic Scrolling Instructions"**

Today you are going search for information using an electronic version of the book. Before you begin, I would like to briefly describe the search tools which are available within the Electronic Scrolling Format. First, as in any book, you can search by using the table of contents. You can move to any section of the book by clicking on the table of contents icon or pulldown menu(Show). The table of contents contains folders of the different chapters. By clicking on a folder that contains a 'plus' symbol, you can view the subtopics of any chapter. Clicking on any one of the subtopics, the tool will take you to the appropriate page. If a 'plus' symbol does not appear on the folder, no additional subtopics will appear (i.e. you are at the lowest level).

A second way to search for information is using the index. The index can be accessed by clicking on the icon on the toolbar or by a pulldown menu on the menu bar (Show). The index provides keywords that are used throughout the text. By double-clicking on a keyword (which contains a page number) or by clicking once on the keyword and pressing the view button to the right of the index listbox, the search tool will take you to the appropriate section of the text. Since a particular keyword might have

multiple page numbers associated with it, the same keyword with different page numbers will presented on different lines.

A third way to search for information us by simply browsing the text. This can be done by manipulating the vertical scrollbar and dragging the square (within the scrollbar) up or down or by clicking on the up or down arrows. This will allow you to move forward and backward through the document. In order to help you navigate through the book, the page number that you are currently on will be displayed at the bottom right of the interface.

There are some features on the interface which will not operational. Since this tool is relatively new, some of the features will not be functional. These features are not broken, they are simply not functional.

2. **Instructions for User Exploration**

Now I will give you approximately 10 Minutes to explore the basic features of the Electronic Scrolling interface on your own and become familiar with the book. Remember that your task is to get acquainted with the interface and its features. You may utilize any feature of book you wish.

USER EXPLORES FOR 10 MINUTES

3. **User Accomplishes Basic Tasks using "Book"**

Now that you have had a chance to explore the Electronic Scrolling interface and utilize some of its features, I would like you to accomplish a series of basic tasks to find information about how to search for information about completing your federal income tax form. For that purpose, we are going to utilize this book in its electronic format. Please go into the book.

(FIRST SESSION ONLY)

This procedure for proceeding through the search tasks will be the same. First, read the search question or query out loud. Then announce when you are ready to begin the search tasks by saying out loud "Ready to Begin". At this time I will initiate the timer and you will have ten minutes to search for the answers. You are then to search until you found the answer.

When you think you have found the answer, I would like you to announce "I have found the answer" or words to that effect, and then physically point to the beginning of the passage that contains the first sentence of the correct answer. I will then ask you to tell me the answer. If I determine that the answer is correct, then I will stop the timer and ask you to read the next search question. If I determine that the answer is correct then I will tell you to continue looking until I tell you to stop. You are to continue searching for the correct answer or the ten minutes have expired.

This procedure will be repeated for each task. Read query out loud, announce when you are ready to being, search for the answer, announce that you found the answer, and then I will ask you for the answer.

By the two of us working together I can better understand how individuals utilize this tool.

(Second and Third Sessions)

You are familiar with the procedures by now but if you have any questions, please don't hesitate to ask.

4. User Completes Electronic Scrolling General Questionnaire

Now I would like you to complete this general questionnaire regarding your subjective impression of Electronic Scrolling interface. In making your evaluation I would like you to focus on the electronic tool, its features, and to the specific contents of the "Book," per se.

5. Session Debrief

Now I would like you to answer some open ended questions regarding this interface.

(CONDUCT DEBRIEF)

APPENDIX K

Electronic Page Flipping Instructions

1. "Introduction to Electronic Page Flipping Instructions'

Today you are going search for information using an electronic version of the book. Before we begin, I would like to briefly describe the search tools which are available within the Electronic Page Flipping Format. First, as in any book, you can search by using the table of contents. You can move to any section of the book by clicking on the table of contents icon or pulldown menu. The table of contents contains folders of the different chapters. By clicking on a folder that contains a 'plus' symbol, you can view the subtopics of any chapter. Clicking on any one of the subtopics, the tool will take you to the appropriate page. If a 'plus' symbol does not appear on the folder, no additional subtopics will appear (i.e. you are at the lowest level).

A second way to search for information is using the index. The index can be accessed by clicking on the icon on the toolbar or by a pulldown menu on the menu bar. The index provides keywords that are used throughout the text. By double-clicking on a keyword (which contains a page number) or by clicking once on the keyword and pressing the view button to the right of the index listbox, the search tool will take you to the appropriate section of the text. Since a particular keyword might have multiple page

numbers associated with it, the same keyword with different page numbers will presented on different lines.

A third way to search for information us by simply browing the text. This can be done in two ways. First, the interface contains a backward and forward button on the left-hand side on the screen. These buttons operate in much the same way as Netscape buttons. By clicking on the forward button, the next page appears on the screen. By clicking on the backward button, the previous page appears on the screen. Another way of browing through the text is to use the horizontal scrollbar at the bottom of the interface. This allows you to move quickly through the pages by holding on the square box within the horizontal scrollbar and dragging it across the scrollbar. You can see the page numbers at the right-hand side of the interface change as you move utilize the scrollbar. A second way to use the horizontal scrollbar is by clicking on the left and right arrows on the sides of the scrollbar. This will either increment or decrement the page that you are currently viewing by one.

(For page flipping long page: In order to see all of the information on the page, a vertical scrollbar is provided for your convenience. By moving the square within the vertical scrollbar wither up or down, the entire page of the document can be viewed. The forward and backward buttons allow for easy navigation from page to page.)

2. Instructions for User Exploration

Now I will give you approximately 10 Minutes to explore the basic features of the electronic page flipping interface on your own. Remember that your task is to get acquainted with the interface and its features. You may utilize any feature of book you wish. There are some features on the interface which will not operational. Since this tool

is relatively new, some of the features will not be functional. These features are not broken, they are simply not functional.

In order to help you navigate through the book, the page number that you are currently on will be displayed at the bottom right of the interface.

USER EXPLORES FOR 10 MINUTES

3. User Accomplishes Basic Tasks using "Book"

(FIRST SESSION ONLY)

Now that you have had a chance to explore the Electronic Page Flipping interface and utilize some of its features, I would like you to accomplish a series of basic tasks to find information about how to search for information about completing your federal income tax form. For that purpose, we are going to utilize this book in its electronic format. Please go into the book.

(FIRST SESSION ONLY)

This procedure for proceeding through the search tasks will be the same. First, read the search question or query out loud. Then announce when you are ready to begin the search tasks by saying out loud "Ready to Begin". At this time I will initiate the timer and you will have ten minutes to search for the answers. You are then to search until you found the answer.

When you think you have found the answer, I would like you to announce "I have found the answer" or words to that effect, and then physically point to the beginning of the passage that contains the first sentence of the correct answer. I will then ask you to tell me the answer. If I determine that the answer is correct, then I will stop the timer and ask you to read the next search question. If I determine that the answer is correct then I

will tell you to continue looking until I tell you to stop. You are to continue searching for the correct answer or the ten minutes have expired.

This procedure will be repeated for each task. Read query out loud, announce when you are ready to being, search for the answer, announce that you found the answer, and then I will ask you for the answer.

By the two of us working together I can better understand how individuals utilize this tool.

(Second and Third Sessions)

You are familiar with the procedures by now but if you have any questions, please don't hesitate to ask.

(Second and Third Sessions)

Repeat the instructions for the finding the answers.

5. User Completes Electronic Page Flipping General Questionnaire

Now I would like you to complete this general questionnaire regarding your subjective impression of Electronic Page Flipping interface. In making your evaluation I would like you to focus on the electronic tool, its features, and to the specific contents of the "Book," per se.

5. Session Debrief

Now I would like you to answer some open ended questions regarding the electronic page flipping tool.

(CONDUCT DEBRIEF)

APPENDIX L

Debriefing Sheet

In the library of the future, books will be presented on the computer. The move to the electronic format precipitated by the lack of storage space in a library and increased cost of printing hardcopy (i.e., paper) books. With libraries and bookstores beginning to put hard copy book information in electronic format, an attempt must be made to understand the difference in the processing of information in its hardcopy and electronic formats.

The present study seeks to understand why information may be found faster in hardcopy format (book) as opposed to the electronic format of the same document. The focus of the study is on the differences in the amount of information available to the user at any one point in time and the failure of the electronic format to the use the book metaphor.

The impetus for the present experiment was a combined usability and marketing assessment of a new electronic library system conducted by Biers et al (1995). To determine the value of a product to the end-user and to identify problems in the use of that product, Biers et al. conducted a comparative usability test in which beginner and intermediate computer users utilized a hardcopy book, an electronic version of the same book and WinHelp (a Windows 95 onscreen help guide) to find information about

Windows 95. The users engaged in a series of search tasks in which they utilized each of the three tools over three sessions.

It was expected that the electronic search tool would aid more readily in finding the information because of the sophistication of the search tool. However, the major findings indicated that there were no differences between the electronic book and hardcopy versions of the book in terms of percent correct. Surprisingly, however, it took the users longer to find the information with the electronic book than the hardcopy documentation. This unexpected finding forms the basis for the present study.

One possible reason that subjects found the information in the hardcopy documentation faster is that more information was available to the user at any one point in time. With the Hardcopy documentation, the book had two pages visible, and the number of lines shown at any one page in the book was more than what could fit on a computer screen. The person could readily scan more information, since the book was back-to-back, two pages at a time. However, with the electronic book, there was a limitation to the number of lines that could be shown on the screen. Also, as the user scrolled down, what he/she has initially seen, disappeared.

Another possible reason for the electronic book and hardcopy difference is that electronic version of the book did not directly conform to a book metaphor with which users have been familiar since an early age (i.e. well-learned habits). In the hardcopy documentation, there are three essential components to this book metaphor. First, the hardcopy documentation is organized in pages, side by side, and one scans for information page to page, horizontally from left to right. Secondly, the text is in a fixed location. As a consequence of this fixed location, the user's eyes move as he/she scans

for information from top to bottom. Thirdly, user can tell where he/she is by the thickness of the book.

However with electronic documentation, the information is typically not organized in discrete pages and information moves on the screen as one moves the vertical scrollbar. In the electronic version of the book, the text actually moves upward as the user scrolls down. The user may start out by scanning from top to bottom but when he/she gets to the bottom of the page and uses the scrollbar, the gaze then becomes fixed as the text moves. This text movement could lead to eye fatigue. The lack of page numbers and the fact that the text does not appear in a fixed location could only add to the user's confusion.

Dillon (1992) conducted a comprehensive review of the factors that could account for the differences in processing information in hardcopy documentation as opposed to the electronic version of the same document. Among the factors reviewed included... The major conclusion that can be drawn from Dillon's review of the empirical studies is that these factors alone do not account for differences in the processing of information using these alternative media. However, most of the studies reviewed by Dillon used short text passages rather that an entire book. Secondly, the focus of the past studies was on reading speed and reading comprehension and not the time it takes to search for information. The present study seeks to determine if two variables not previously researched are responsible for information processing differences in hardcopy and electronic formats of the same document, and to extend the research literature to full-length text (books) and to search tasks.

The purpose of the present study was to investigate the factors responsible for differences in searching for information in hardcopy and electronic formats. The two reasons being investigated are the amount of text which is available for the user at any one given point in time and the failure of electronic version to conform to the book metaphor.

(CONDUCT DEBRIEF)

REFERENCES

Biers, D., Anthony C., Blinn, B., Buckley, B., Chambers, W., Salyers, M. (1995) Marketing Assessment of VIRDOX Software, (Report No. 9495-150). Dayton, OH: Center for Business and Economic Research.

Dillon, A., Richardson, J., and McKnight, C. 1990b, The effect of display size and paragraph splitting on reading lengthy text from screen, Behaviour and Information Technology, 9, 215-227.

Dillon, A.,1992, Reading from paper versus screens: a critical review of the empirical literature, Ergonomics, 35, 1297-1326.

Duchnicky, R.L. and Kolers, P.A. 1983. Readability of text scrolled on a visual display terminal as a function of window size, Human Factors, 25, 683-692.

Gould, J.D. and Grischowsky, N. 1984. Doing the same work with hardcopy and cathode ray tube (CRT) computer terminals. Human Factors, 26, 323-337.

Gould, J.D. Alfaro, L., Barnes, V., Finn, R., Grischowsky, N. and Minuto, S. 1987a, Reading is slower from CRT displays than from paper: Attempts to isolate a single variable explanation, Human Factors, 29, 269-299.

Gould, J.D. Alfaro, L., Finn, R., Haupt, B., Grischowsky, N. and Minuto, S. 1987b. Reading from CRT displays can be as fast as paper, Human Factors, 26, 497-517.

Kruk, R.S. and Muter, P. 1984. Reading continuous text on video screens, Human Factors, 26, 339-345.

Mills, C. B. and Weldon, L.J. 1986. Reading text from computer screens. ACM Computing Surveys, 19, 328-358.

Richardson, J., Dillon, A., and McKnight, C. 1989. The effect of window size on reading and manipulating electronic text, in E. Megaw (ed.) Contemporary Ergonomics, 1989 (Taylor and Francis, London).

Scharwtz, E., Beldie, I., and Pastoor, S. 1983. A comparison of paging and scrolling for changing screen contents by inexperienced users, Human Factors, 25, 279-282.